

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:03:03 ; Search time 1026 Seconds

(without alignments)
7337.044 Million cell updates/sec

Title: US-10-009-570-1

Perfect score: 1772

Sequence: 1 acttagataataagaagtaag.....ctcttccttcagtgcaag 1772

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1772	100.0	1772	4	Aaf24785 Nucleotid
2	1772	100.0	1803	4	Aaf24787 Nucleotid
3	1772	100.0	2668	4	Aaf24791 Complete
4	1686	95.1	1686	4	Aaf24790 Nucleotid
5	660	37.2	2000	7	Ada73179 Rice gene
6	86	4.9	86	4	Aaf24788 Nucleotid
7	45	2.5	2000	7	Ada72764 Rice gene
8	34	1.9	2000	7	Ada72806 Rice gene
9	32	1.8	2000	7	Ada72521 Rice gene
10	30	1.7	2000	7	Ada73205 Rice gene
11	28	1.6	2000	7	Ada73342 Rice gene
12	28	1.6	5579	6	AAL46958 Rice lesi
13	26	1.5	2000	7	Ada73166 Rice gene
14	26	1.5	2000	7	Ada73198 Rice gene
15	24	1.4	676	4	Aaf81474 Corn prom
16	24	1.4	2000	7	Ada71850 Rice gene
17	24	1.4	2199	9	Adc08382 Rice DNA
18	23	1.3	1999	9	Adc08516 Rice DNA
19	23	1.3	2000	7	Ada71502 Rice gene
20	23	1.3	2000	7	Ada72983 Rice gene
21	23	1.3	2000	9	Adc08500 Rice DNA
22	23	1.3	5579	6	AAL46958 Rice lesi
23	23	1.3	6533	6	ABK28384 DNA trans

24	23	1.3	6544	4	AAS45340	Chemical
25	23	1.3	6544	6	ABL32646	Human imm
26	23	1.3	7442	4	AAS46686	Tumour su
27	23	1.3	9052	6	ABL32145	Human imm
28	23	1.3	9267	6	ABL33853	Human imm
29	23	1.3	9817	6	ABL33368	Human imm
30	23	1.3	90541	6	ABSS2847	Human SR
31	23	1.3	113033	7	AAL54213	SR protei
32	22	1.2	22	4	Aaf24794	PCR prime
33	22	1.2	1999	9	ADC08514	Rice DNA
34	22	1.2	2000	7	ADA72153	Rice gene
35	22	1.2	2000	7	ADA72089	Rice gene
36	22	1.2	2000	7	ADA73195	Rice gene
37	22	1.2	2000	7	ADA73343	Rice gene
38	22	1.2	2000	7	ADA73432	Rice gene
39	22	1.2	2000	7	ADA71502	Rice gene
40	22	1.2	2000	7	ADA73460	Rice gene
41	22	1.2	2000	7	ADA71850	Rice gene
42	22	1.2	2000	7	ADA72700	Rice gene
43	22	1.2	6061	6	ABL70591	Chemical
44	22	1.2	6061	6	AAS61302	Human gen
45	22	1.2	6185	6	ABL32412	Human imm

ALIGNMENTS

RESULT 1

AAF24785
ID AAF24785 standard; DNA; 1772 BP.
XX
AC AAF24785;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of a rice sucrose synthase 3 (RSus3) promoter.
XX
KW Rice; sucrose synthase 3 promoter; RSus3 promoter; endosperm expression;
KW transgenic plant; carbohydrate metabolism; sucrose metabolism; ss.
XX
OS Oryza sativa.
XX
PN WO200078975-A2.
XX
PD 28-DEC-2000.
XX
PF 15-JUN-2000; 2000WO-GB002641.
XX
PR 17-JUN-1999; 99GB-00014210.
XX
PA (DANI-) DANISCO AS.
XX
PI Donaldson IA, Rasmussen TB;
XX
DR WPI; 2001-071398/08.
XX
PT Rice sucrose synthase 3 promoter obtainable from plant genus Oryza,
PT useful for expressing nucleotide sequence of interest in specific tissue
or cell type e.g. endosperm.
XX
PS Claim 1; Page 142; 144pp; English.
XX
CC The present sequence represents a rice sucrose synthase 3 (RSus3)
CC promoter. The RSus3 promoter has low homology with the RSus1 and RSus2
CC promoters (7.7% and 4.6%, respectively). In addition, apart from
CC conserved TATA box and intron splice sites, they have no motifs in
CC common. The promoter can be used to prepare a product of interest,
CC especially in the endosperm of a transgenic plant. Preferably, the
CC promoter is used to cause expression of a nucleotide of interest that can
CC affect carbohydrate metabolism, such as sucrose metabolism, in a plant
tissue
XX
SQ Sequence 1772 BP; 540 A; 368 C; 374 G; 490 T; 0 U; 0 Other;

961 GCATGATGGCTCTGTGAAACAAACAAAGGTTACTGCTAATGACATGCTGCTGTAGCT 1020
 961 GCATGATGGCTCTGTGAAACAAACAAAGGTTACTGCTAATGACATGCTGCTGTAGCT 1020
 1021 AGTTAGCAGATGCAAGGCCCATGCAATGCAATGCTATGCAACAAAGTATAGTACACGCA 1080
 1021 AGTTAGCAGATGCAAGGCCCATGCAATGCAATGCTATGCAACAAAGTATAGTACACGCA 1080
 1081 TGTATGTTAGCAGCTAACTAATCTATCAGCAGAGCGCAGAGCTGCTGCAATGCTGTAT 1140
 1081 TGTATGTTAGCAGCTAACTAATCTATCAGCAGAGCGCAGAGCTGCTGCAATGCTGTAT 1140
 1141 GCATCTCTCCAGTAATCTAGTGTAAATTTTACCCAAAGCGTTGCTCATATGACACAGT 1200
 1141 GCATCTCTCCAGTAATCTAGTGTAAATTTTACCCAAAGCGTTGCTCATATGACACAGT 1200
 1201 AATTAGTAATATACCAAGGTTCAACATCCCGTTACCTGACCAATATCTACTCAGCAATG 1260
 1201 AATTAGTAATATACCAAGGTTCAACATCCCGTTACCTGACCAATATCTACTCAGCAATG 1260
 1261 GTATCTCTGTTTTCGTTAAACCGTTGGTAAACCGTAAACCGTAAACCGTAAACCGTAA 1320
 1261 GTATCTCTGTTTTCGTTAAACCGTTGGTAAACCGTAAACCGTAAACCGTAAACCGTAA 1320
 1321 AATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1380
 1321 AATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1380
 1381 AGCTCTCTGGAAGGACGGTTTGGTGGGATTTGGAACCTGTTGAAACCTGTTACTGCACTTCAT 1440
 1381 AGCTCTCTGGAAGGACGGTTTGGTGGGATTTGGAACCTGTTGAAACCTGTTACTGCACTTCAT 1440
 1441 TTGTAACAGATATTAGTCAACAGACAAATGCAACCGCATTTTCTGTTTACCGGCA 1500
 1441 TTGTAACAGATATTAGTCAACAGACAAATGCAACCGCATTTTCTGTTTACCGGCA 1500
 1501 GCTGAAGCTTTTACGATCCCATACCGCGTTGCTGCAACCTGCTGCAACCTGCTGCAACCTG 1560
 1501 GCTGAAGCTTTTACGATCCCATACCGCGTTGCTGCAACCTGCTGCAACCTGCTGCAACCTG 1560
 1561 GAAACAGGTTGCTCAATTTTGTGGTGGAAAGCCAAAGTAAAGTAAACAGAGATGGAAGATAGT 1620
 1561 GAAACAGGTTGCTCAATTTTGTGGTGGAAAGCCAAAGTAAAGTAAACAGAGATGGAAGATAGT 1620
 1621 GAGGACAGGGAGTGAGGAGGAGGACATGCGCCACGCTCCCTGCAATTTTCTGCTGTA 1680
 1621 GAGGACAGGGAGTGAGGAGGAGGACATGCGCCACGCTCCCTGCAATTTTCTGCTGTA 1680
 1681 TAAATA 1686
 1681 TAAATA 1686

RESULT 5
 ADA73179
 ID ADA73179 standard; DNA; 2000 BP.
 XX
 AC ADA73179;
 XX
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX
 DE Rice gene, SEQ ID 6505.
 XX
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX
 XX Oryza sativa.
 OS
 XX
 XX WO200300898-A1.
 XX
 XX
 PD 03-JAN-2003.
 XX

22-JUN-2001; 2001WO-IB0011105.
 22-JUN-2001; 2001WO-IB0011105.
 (SYGN) SYNGENTA PARTICIPATIONS AG.
 Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 WPI; 2003-175290/17.
 Identifying at least one gene involved in plant resistance or response to
 pathogenic infection for conferring resistance or tolerance to a plant to
 bacterial, fungal or viral infection by determining or detecting plant
 gene expression.
 Claim 27; SEQ ID NO 6505; 899pp; English.
 The present invention relates to a method (M1) for identifying genes
 involved in plant resistance or response to pathogenic infection. M1
 comprises identifying a gene whose expression is significantly altered in
 the incompatible interaction of plant gene expression relative to
 the expression of the gene in an uninfected plant, in a mutant plant that
 does not express a gene associated with response to pathogenic infection,
 or in a corresponding incompatible or compatible interaction. (M1) is
 useful for conferring resistance to resistance or tolerance to a plant to
 bacterial, fungal or viral infection. The present sequence was used to
 illustrate the invention.
 Sequence 2000 BP; 520 A; 394 C; 427 G; 658 T; 0 U; 1 Other;
 Query Match 37.2%; Score 660; DB 7; Length 2000;
 Best Local Similarity 99.4%; Pred. No. 3.8e-301;
 Matches 1090; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 669 CGCTAATAGTCAGATATAAAATTTTATCTATAGATCAATTAATCAATTTTGTGCT 728
 1 CGCTAATAGTCAGATATAAAATTTTATCTATAGATCAATTAATCAATTTTGTGCT 60
 729 TCGTTTCAATTTTCTACCACTTATCAACCATAGCTCACTCAATTTGCAATAAAGTT 788
 61 TCGTTTCAATTTTCTACCACTTATCAACCATAGCTCACTCAATTTGCAATAAAGTT 120
 789 ACTAAACGACATCGCTCATCAACACCCCAACCTCCGATCGGTCCTCTCGACACGA 848
 121 ACTAAACGACATCGCTCATCAACACCCCAACCTCCGATCGGTCCTCTCGACACGA 180
 849 GTTTAGCACTTGTGCAACATATATGCTGCGATGACATCTACTGATGCGCATGCAAT 908
 181 GTTTAGCACTTGTGCAACATATATGCTGCGATGACATCTACTGATGCGCATGCAAT 240
 909 TTTAGCTTCTGTTTCATGAGCTTCCAAACGACAGAGGCTGAGCAGCAGCATGCAAT 968
 241 TTTAGCTTCTGTTTCATGAGCTTCCAAACGACAGAGGCTGAGCAGCAGCATGCAAT 300
 969 GGCTCTTGTGAAACCAAAAAGGTTACTGTTAAATGACATGCTGTTAGTAGCA 1028
 301 GGCTCTTGTGAAACCAAAAAGGTTACTGTTAAATGACATGCTGTTAGTAGCA 360
 1029 GAATGCAAGCCCATGTCATGCAATGCTTATGCAACAGTATAGTACCAAGCATGATGT 1088
 361 GAATGCAAGCCCATGTCATGCAATGCTTATGCAACAGTATAGTACCAAGCATGATGT 420
 1089 AGCAGCTAACTAATCTATCAGCAGAGGCAAGCTCGTGTGATGTTGATGCACTTCT 1148
 421 AGCAGCTAACTAATCTATCAGCAGAGGCAAGCTCGTGTGATGTTGATGCACTTCT 480
 1149 CTCAGTAATCTAGTGGTAAATTTTCAACCAAGCGTTGCTCATATGACAGTAATAGTA 1208
 481 CTCAGTAATCTAGTGGTAAATTTTCAACCAAGCGTTGCTCATATGACAGTAATAGTA 540
 1209 AATTTACCAAGTTTCAATCCGTTACCTGACCAATATCTACTCAGCAATGTTATCTCT 1268

[illegible]

XX	CC	The present sequence represents a rice sucrose synthase 3 (RSus3)	
CC	CC	promoter fragment. The RSus3 promoter has low homology with the RSus1 and	
CC	CC	RSus2 promoters (7.7% and 4.6%, respectively). In addition, apart from	
CC	CC	conserved TATA box and intron splice sites, they have no motifs in	
CC	CC	common. The promoter can be used to prepare a product of interest,	
CC	CC	especially in the endosperm of a transgenic plant. Preferably, the	
CC	CC	promoter is used to cause expression of a nucleotide of interest that can	
CC	CC	affect carbohydrate metabolism, such as sucrose metabolism, in a plant	
CC	CC	tissue	
XX	CC		
SQ	Sequence	86 BP; 13 A; 29 C; 16 G; 28 T; 0 U; 0 Other;	
XX	Query Match	4.9%; Score 86; DB 4; Length 86;	
XX	Best Local Similarity	100.0%; Pred. No. 5e-30;	
XX	Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX	CC		
QY	1687	CAGTGATGATCGCTCTCCAGATCCATCGGTTCTCTGCTGTTCATCCATAGT	1746
Db	1	CAGTGATGATCGCTCTCCAGATCCATCGGTTCTCTGCTGTTCATCCATAGT	60
XX	CC		
QY	1747	TTCTCTCTCTCTCCCTCAGTGCAG 1772	
XX	CC		
Db	61	TTCTCTCTCTCTCCCTCAGTGCAG 86	
XX	CC		
RESULT 7	ADA72764/C		
ID	ADA72764	standard; DNA; 2000 BP.	
XX	AC	ADA72764;	
XX	CC		
DT	20-NOV-2003	(first entry)	
XX	CC		
DE	Rice gene, SEQ ID 6089.		
XX	CC		
KW	Plant; bacterial infection; fungal infection; viral infection; rice;		
XX	gene; ds.		
XX	Oryza sativa.		
XX	OS		
XX	WO2003000898-A1.		
PN	22-JUN-2001; 2001WO-IB001105.		
XX	CC		
XX	2003-2003.		
PD	22-JUN-2001; 2001WO-IB001105.		
XX	CC		
PF	22-JUN-2001; 2001WO-IB001105.		
XX	CC		
PR	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX	CC		
PA	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;		
XX	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;		
PI	WPI; 2003-1752290/17.		
XX	CC		
DR	Identifying at least one gene involved in plant resistance or response to		
XX	pathogenic infection for conferring resistance or tolerance to a plant to		
PT	bacterial, fungal or viral infection by determining or detecting plant		
XX	gene expression.		
XX	CC		
PS	Claim 27; SEQ ID NO 6089; 899pp; English.		
XX	CC		
CC	The present invention relates to a method (M1) for identifying genes		
CC	involved in plant resistance or response to pathogenic infection. M1		
CC	comprises identifying a gene whose expression is significantly altered in		
CC	the incompatible interaction of plant gene expression relative to		
CC	expression of the gene in an uninfected plant, in a mutant plant that		
CC	does not express a gene associated with response to pathogenic infection,		
CC	or in a corresponding incompatible or compatible interaction. (M1) is		
CC	useful for conferring resistance to resistance or tolerance to a plant to		
CC	bacterial, fungal or viral infection. The present sequence was used to		
CC	illustrate the invention.		

XX 03-JAN-2003.
 XX PD
 XX PF
 XX 22-JUN-2001; 2001WO-IB001105.
 XX PR
 XX 22-JUN-2001; 2001WO-IB001105.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 XX
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 XX Claim 27; SEQ ID NO 6531; 899pp; English.
 XX
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 XX Sequence 2000 BP; 613 A; 404 C; 328 G; 655 T; 0 U; 0 Other;
 XX
 XX Query Match 1.7%; Score 30; DB 7; Length 2000;
 XX Best Local Similarity 100.0%; Pred. No. 0.0013;
 XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 47 AAAAAAAAAAAGACGAGTGTCAAC 76
 XX ||||||||||||||||||||||||||||
 XX 367 AAAAAAAAAAAGACGAGTGTCAAC 338
 XX
 XX RESULT 11
 XX ADA73342
 XX ID ADA73342 standard; DNA; 2000 BP.
 XX AC ADA73342;
 XX
 XX 20-NOV-2003 (first entry)
 XX DE Rice gene, SEQ ID 6668.
 XX
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 XX gene; ds.
 XX
 XX Oryza sativa.
 XX
 XX WO2003000898-A1.
 XX
 XX 03-JAN-2003.
 XX
 XX 22-JUN-2001; 2001WO-IB001105.
 XX
 XX 22-JUN-2001; 2001WO-IB001105.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 XX
 XX

PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 XX Claim 27; SEQ ID NO 6668; 899pp; English.
 XX
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 XX Sequence 2000 BP; 638 A; 344 C; 433 G; 583 T; 0 U; 2 Other;
 XX
 XX Query Match 1.6%; Score 28; DB 7; Length 2000;
 XX Best Local Similarity 100.0%; Pred. No. 0.011;
 XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 29 AAAAAATAATAATATCCAAATTTTTT 56
 XX ||||||||||||||||||||||||||||
 XX 292 AAAAAATAATAATATCCAAATTTTTT 319
 XX
 XX RESULT 12
 XX AAL46958/c
 XX ID AAL46958 standard; cDNA; 5579 BP.
 XX AC AAL46958;
 XX
 XX 30-AUG-2002 (first entry)
 XX DE Rice lesion inhibitor protein Spl7 coding sequence.
 XX
 XX Rice; lesion formation inhibition; heat stress; agriculture; Spl7;
 XX transgenic; plant; horticulture; gene; ss.
 XX
 XX Oryza sativa.
 XX
 XX Key Location/Qualifiers
 XX exon 3177..3947
 XX /tag= b
 XX /number= 1
 XX CDS 3711..5327
 XX /tag= a
 XX /product= "Spl7"
 XX intron 3948..4184
 XX /tag= c
 XX /number= 1
 XX exon 4185..5327
 XX /tag= d
 XX /number= 2
 XX
 XX WO200233092-A1.
 XX
 XX 25-APR-2002.
 XX
 XX 18-OCT-2001; 2001WO-JF009153.
 XX
 XX 18-OCT-2000; 2000JP-00318557.
 XX (NAAG-) NAT INST AGROBIOLOGICAL SCI.
 XX Yano M, Yamanouchi U;
 XX WPI; 2002-372312/40.
 XX P-PSDB; AA017798.
 XX
 XX

PT Rice-originated gene, Spi17, that inhibits lesion formation and is
PT applicable in improving heat stress of plants thus leading to prevention
PT of lesion formation, for developing new breeds of plants for agriculture
PT and horticulture.

XX Claim 1; Page 30-40; 53pp; Japanese.

XX The present invention provides the protein and coding sequences of rice
CC lesion formation inhibitor Spi17. The protein improves the heat stress of
CC the plant, and can be used in the development of new breeds of plants for
CC agriculture and horticulture. The present sequence is the coding sequence
CC of the invention

XX Sequence 5579 BP; 1453 A; 1272 C; 1235 G; 1619 T; 0 U; 0 Other;

Query Match 1.6%; Score 28; DB 6; Length 5579;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 AATAAGACGAGTGGTCAAAACAGTACAAG 84

DB 2210 AATAAGACGAGTGGTCAAAACAGTACAAG 2183

RESULT 13

ADA73166/C
ID ADA73166 standard; DNA; 2000 BP.

XX AC ADA73166;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 6492.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-Al.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX Claim 27; SEQ ID NO 6492; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the incompatible interaction of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 2000 BP; 492 A; 534 C; 468 G; 504 T; 0 U; 2 Other;

Query Match 1.5%; Score 26; DB 7; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTTAGATATAAAGTAAGTCACAA 26

DB 1060 ACTTTAGATATAAAGTAAGTCACAA 1035

RESULT 14

ADA73198
ID ADA73198 standard; DNA; 2000 BP.

XX AC ADA73198;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 6524.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-Al.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX Claim 27; SEQ ID NO 6524; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the incompatible interaction of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 2000 BP; 633 A; 407 C; 355 G; 604 T; 0 U; 1 Other;

Query Match 1.5%; Score 26; DB 7; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTTAGATATAAAGTAAGTCACAA 26

DB 108 ACTTTAGATATAAAGTAAGTCACAA 133

RESULT 15

AAF81474/C
ID AAF81474 standard; DNA; 676 BP.

XX

AC AAF81474;
XX
DT 08-JUN-2001 (first entry)
XX
DE Corn promoter clone #700345819.
XX
KW Corn; promoter; transgenic plant; herbicide resistance; ds.
XX
OS Zea mays.
XX
FN WO200119976-A2.
XX
PD 22-MAR-2001.
XX
PF 13-SEP-2000; 2000WO-US025078.
XX
PR 16-SEP-1999; 99US-0154182P.
XX
PA (MONS) MONSANTO CO.
XX
PI Anderson HM, Chay CA, Chen G, Conner TW;
XX
DR WPI; 2001-244796/25.
XX
PT Novel promoter nucleic acid sequences useful for regulating heterologous
PT gene expression in plants, comprising regulatory sequences located
PT upstream to plant DNA structural coding sequences.
XX
PS Claim 1; Page 98; 101pp; English.
XX
CC The present invention relates to novel corn promoter sequences (see
CC AAF81456-AAF81478). The promoter sequences are useful for conferring
CC expression of a second polynucleotide molecule in a transgenic plant
CC tissue. In addition, the promoter sequences are useful for providing
CC plants with herbicide resistance. The promoter sequences are suitable for
CC selectively modulating expression of any operatively linked gene and
CC provide additional regulatory element diversity in a plant expression
CC vector in gene stacking approaches. The present sequence is one such corn
CC promoter sequence isolated in the present invention
XX
SQ Sequence 676 BP; 210 A; 152 C; 119 G; 195 T; 0 U; 0 Other;

Query Match 1.4%; Score 24; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 TTTTAAATAAGACGAGTGTCAA 74
|||||
Db 226 TTTTAAATAAGACGAGTGTCAA 203
|||||

Search completed: October 2, 2004, 10:13:35
Job time : 1029 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 08:40:19 ; Search time 153 Seconds
(without alignments)
6427.274 Million cell updates/sec

Title: US-10-009-570-1

Perfect score: 1772

Sequence: 1 actttagataataagaag.....ctcttctccttcagtcgaag 1772

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24	1.4	676	4	US-09-665-189A-71
C 2	23	1.3	90541	4	US-09-759-359A-3
C 3	23	1.3	193303	4	US-09-497-855A-37
C 4	23	1.3	193303	4	US-09-497-855A-44
C 5	20	1.1	177	4	US-09-313-294A-292
C 6	20	1.1	888	4	US-09-134-001C-13
C 7	20	1.1	1462	2	US-08-464-517-30
C 8	20	1.1	1462	2	US-08-246-361A-30
C 9	20	1.1	1462	3	US-08-463-772-30
C 10	20	1.1	1462	5	PCT-US93-05000-30
C 11	19	1.1	531	3	US-08-642-274D-39
C 12	19	1.1	531	3	US-08-952-014C-39
C 13	19	1.1	573	4	US-09-134-001C-301
C 14	19	1.1	851	4	US-09-495-050A-138
C 15	19	1.1	2095	4	US-09-227-357-31
C 16	19	1.1	6113	4	US-10-204-708-14
C 17	19	1.1	6574	4	US-09-321-017B-1097
C 18	19	1.1	6669	4	US-10-204-708-6
C 19	19	1.1	6678	3	US-08-816-617A-1
C 20	19	1.1	11049	4	US-10-204-708-23
C 21	19	1.1	25464	4	US-09-326-480A-4
C 22	19	1.1	70000	4	US-09-851-896-3
C 23	19	1.1	392000	4	US-10-027-983-11
C 24	19	1.1	1830121	4	US-09-557-884-1
C 25	19	1.1	1830121	4	US-09-643-990A-1
C 26	18	1.0	405	4	US-09-621-976-15602
C 27	18	1.0	562	5	PCT-US92-03993-6

C 28	18	1.0	603	4	US-09-322-478-35	Sequence 35, Appli
C 29	18	1.0	630	4	US-09-342-681C-5	Sequence 5, Appli
C 30	18	1.0	705	4	US-09-134-000C-1679	Sequence 1679, Ap
C 31	18	1.0	1587	4	US-09-543-681A-1750	Sequence 1750, Ap
C 32	18	1.0	1881	4	US-09-976-594-764	Sequence 764, App
C 33	18	1.0	2230	4	US-09-620-312D-282	Sequence 282, App
C 34	18	1.0	2502	1	US-08-062-472B-2	Sequence 2, Appli
C 35	18	1.0	2951	4	US-09-177-419C-3	Sequence 3, Appli
C 36	18	1.0	3600	4	US-09-134-000C-2137	Sequence 2137, Ap
C 37	18	1.0	4088	2	US-08-317-310A-1	Sequence 1, Appli
C 38	18	1.0	4088	5	PCT-US95-13041-1	Sequence 1, Appli
C 39	18	1.0	4609	4	US-09-322-478-21	Sequence 21, Appli
C 40	18	1.0	4711	1	US-08-414-926A-1	Sequence 1, Appli
C 41	18	1.0	4711	2	US-08-926-922-1	Sequence 1, Appli
C 42	18	1.0	4711	3	US-09-253-682-1	Sequence 1, Appli
C 43	18	1.0	4711	3	US-09-527-657-1	Sequence 1, Appli
C 44	18	1.0	4711	4	US-09-892-100-1	Sequence 1, Appli
C 45	18	1.0	5152	4	US-10-204-708-48	Sequence 48, Appli

ALIGNMENTS

RESULT 1

US-09-665-189A-71/c
; Sequence 71, Application US/09665189A
; Patent No. 6645765
; GENERAL INFORMATION:
; APPLICANT: Anderson, Heather
; APPLICANT: Chay, Catherine
; APPLICANT: Chen, Guilan
; APPLICANT: Conner, Timothy
; TITLE OF INVENTION: Plant Regulatory Sequences for Control of Gene Expression
; FILE REFERENCE: 38-211(15674)B
; CURRENT APPLICATION NUMBER: US/09/665,189A
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/665,189
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 71
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Zea mays
US-09-665-189A-71

Query Match 1.4%; Score 24; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	51	TTTTTTAATAAGACGAGTGTCAA	74
DB	226	TTTTTTAATAAGACGAGTGTCAA	203

RESULT 2

US-09-759-359A-3/c
; Sequence 3, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA
; ORGANISM: Human

US-09-759-359A-3

Query Match 1.3%; Score 23; DB 4; Length 90541;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AAGTAAGTCACAGAAAATAAA 36
DB 17920 AAGTAAGTCACAGAAAATAAA 17898

RESULT 3

US-09-497-855A-37
; Sequence 37, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37

Query Match 1.3%; Score 23; DB 4; Length 193303;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1744 AGTTTCCTCTCTCTCTCTCTCTCAG 1766
DB 166834 AGTTTCCTCTCTCTCTCTCTCAG 166856

RESULT 4

US-09-497-855A-44
; Sequence 44, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-44

Query Match 1.3%; Score 23; DB 4; Length 193303;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1744 AGTTTCCTCTCTCTCTCTCTCTCAG 1766
DB 166834 AGTTTCCTCTCTCTCTCTCTCAG 166856

RESULT 5

US-09-313-294A-292/c
; Sequence 292, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 292
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700548929H1
; NAME/KEY: unsure
; LOCATION: 2, 6, 75-93
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-292

Query Match 1.1%; Score 20; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 950 AGCAGCAGCATGCATGCATG 969
DB 59 AGCAGCAGCATGCATGCATG 40

RESULT 6

US-09-134-001C-13/c
; Sequence 13, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 13
; LENGTH: 888
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (505)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-13

Query Match 1.1%; Score 20; DB 4; Length 888;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1285 GTTGTAACCCAGCAAAAT 1304
DB 350 GTTGTAACCCAGCAAAAT 331

RESULT 7

US-08-464-517-30
; Sequence 30, Application US/08464517

```
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1462 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(378..569, 662..1000, 1040..1189,
; LOCATION: 1191..1292, 1292..1324)
; US-08-464-517-30
```

```
Query Match 1.1%; Score 20; DB 2; Length 1462;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 AAAAACAACAAACACAC 177
Db 310 AAAAACAACAAACACAC 329
```

```
RESULT 8
US-08-246-361A-30
; Sequence 30, Application US/08246361A
; Patent No. 598582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1462 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(378..569, 662..1000, 1040..1189,
; LOCATION: 1191..1292, 1292..1324)
; US-08-246-361A-30

Query Match 1.1%; Score 20; DB 2; Length 1462;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 AAAAACAACAAACACAC 177
Db 310 AAAAACAACAAACACAC 329

RESULT 9
US-08-463-772-30
; Sequence 30, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
```

;; FILING DATE: 16-OCT-1992
;; APPLICATION NUMBER: US 07/888,178
;; FILING DATE: 26-MAY-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/701,514
;; FILING DATE: 16-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Matthew P. Vincent
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-004C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-7400
;; TELEFAX: (617) 227-5941
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1462 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: join(378..569, 662..1000, 1040..1189,
;; LOCATION: 1191..1292, 1292..1324)
US-08-463-772-30

Query Match 1.1%; Score 20; DB 3; Length 1462;
Best Local Similarity 100.0%; Pred. No. 9.7; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 158 AAAAACCAAAACAAACACAC 177
Db 310 AAAAACCAAAACAAACACAC 329
|||||

RESULT 10
PCT-US93-05000-30
; Sequence 30, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOXIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1462 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double

;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
PCT-US93-05000-30

Query Match 1.1%; Score 20; DB 5; Length 1462;
Best Local Similarity 100.0%; Pred. No. 9.7; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 158 AAAAACCAAAACAAACACAC 177
Db 310 AAAAACCAAAACAAACACAC 329
|||||

RESULT 11
US-08-642-274D-39
; Sequence 39, Application US/08642274D
; Patent No. 6200749
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
; FILE REFERENCE: 229000033
; CURRENT APPLICATION NUMBER: US/08/642,274D
; CURRENT FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic
US-08-642-274D-39

Query Match 1.1%; Score 19; DB 3; Length 531;
Best Local Similarity 100.0%; Pred. No. 30; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

Qy 1321 AATTTAAATTTAGTTT 1339
Db 105 AATTTAAATTTAGTTT 123
|||||

RESULT 12
US-08-952-014C-39
; Sequence 39, Application US/08952014C
; Patent No. 6265158
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
; TITLE OF INVENTION: GENOMIC ORGANIZATION
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6265158thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952,014C
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00028
; TELECOMMUNICATION INFORMATION:


```
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 531 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-952-014C-39

Query Match
Best Local Similarity 100.0%; Score 19; DB 3; Length 531;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1321 AATTTAAATTTAGTTT 1339
Db 105 AATTTAAATTTAGTTT 123

RESULT 13
US-09-134-001C-301/c
; Sequence 301, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GFC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 301
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-301

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 573;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 710 TAAATCATTTTAGTTGCT 728
Db 99 TAAATCATTTTAGTTGCT 81

RESULT 14
US-09-495-050A-138/c
; Sequence 138, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 138
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; OTHER INFORMATION: Incyte ID No. 6492505 1516908CB1
US-09-495-050A-138

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 851;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 AAAAAACAAACAAACA 174
Db 117 AAAAAACAAACAAACA 99

RESULT 15
US-09-227-357-31/c
; Sequence 31, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
```

```

; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 31
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-31

```

```

Query Match      1.1%; Score 19; DB 4; Length 2095;
Best Local Similarity 100.0%; Pred.No. 28;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 158 AAAAACAAAAAACACACA 176
    |||||
Db 489 AAAAACAAAAAACACACA 471

```

Search completed: October 2, 2004, 15:15:58
Job time : 157 secs

RESULTS 2

US-10-437-963-92628
; Sequence 92628, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 92628
; LENGTH: 2981
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2981)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91091C.1
US-10-437-963-92628

Query Match 2.1%; Score 37; DB 17; Length 2981;
Best Local Similarity 100.0%; Pred. No. 8.5e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1727 GCTCTGTTCCATCCATAGAGTTTCTCTCTCTCTCTCTCTCT 1763
Db 152 GCTCTGTTCCATCCATAGAGTTTCTCTCTCTCTCTCTCTCT 188

RESULT 3
US-10-168-273B-1/c
; Sequence 1, Application US/10168273B
; Publication No. US20040058324A1
; GENERAL INFORMATION:
; APPLICANT: Yano, Masahiro
; APPLICANT: Yamanouchi, Utako
; TITLE OF INVENTION: PLANT LESION FORMATION SUPPRESSING GENE, Sp17 AND USE THEREOF
; FILE REFERENCE: 23572-005 NATL
; CURRENT APPLICATION NUMBER: US/10/168,273B
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/JP01/09153
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: JP 2000-318557
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: genomic DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3711)..(3947)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4185)..(5327)
US-10-168-273B-1

Query Match 1.6%; Score 28; DB 13; Length 5579;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 AATAAGACGAGTGGTCAAAACAGTACAA 84
Db 2210 AATAAGACGAGTGGTCAAAACAGTACAA 2183

RESULT 4
US-10-260-238-1731/c
; Sequence 1731, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Krepas, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1731
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1731

Query Match 1.5%; Score 27; DB 16; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 AATAAGACGAGTGGTCAAAACAGTACAA 83
Db 1064 AATAAGACGAGTGGTCAAAACAGTACAA 1038

RESULT 5
US-10-437-963-27816
; Sequence 27816, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 27816
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure

```
; LOCATION: (1)..(630)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32474C.1
US-10-437-963-27816

Query Match
Best Local Similarity 1.5%; Score 26; DB 17; Length 630;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTTAGATAATAAGTAAGTCAAA 26
Db 428 ACTTTAGATAATAAGTAAGTCAAA 453

RESULT 6
US-10-425-114-718/c
; Sequence 718, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 718
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700087978_FLI
US-10-425-114-718

Query Match
Best Local Similarity 1.4%; Score 25; DB 13; Length 858;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATTTTATTAATAGACGAGTGCTCA 73
Db 655 ATTTTATTAATAGACGAGTGCTCA 631

RESULT 7
US-10-260-238-2163
; Sequence 2163, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Barrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
```

```
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 2163
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-2163

Query Match
Best Local Similarity 1.4%; Score 25; DB 16; Length 2000;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 AAAAAGCTCAAAATTCCTTATATTAT 110
Db 81 AAAAAGCTCAAAATTCCTTATATTAT 105

RESULT 8
US-10-437-963-10006/c
; Sequence 10006, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 10006
; LENGTH: 6064
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_16369C.1
US-10-437-963-10006

Query Match
Best Local Similarity 1.4%; Score 25; DB 17; Length 6064;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CTCAAAATTCCTTATATTATGGGAC 115
Db 3129 CTCAAAATTCCTTATATTATGGGAC 3105

RESULT 9
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MEC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
```

; LOCATION: (379615)
US-10-312-841-2

Query Match 1.4%; Score 25; DB 15; Length 3673778;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1322 ATTATTAATTTAGTTTTTTTAA 1346
Db 729897 ATTTAAATTTAGTTTTTTTAA 729921

RESULT 10

US-10-437-963-87728
; Sequence 87728, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87728
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86647C.1
US-10-437-963-87728

Query Match 1.4%; Score 24; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 GATATAAGTAAGTCACAGAAA 30
Db 309 GATATAAGTAAGTCACAGAAA 332

RESULT 11

US-10-424-599-64432
; Sequence 64432, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 64432
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29194C.1
US-10-424-599-64432

Query Match 1.4%; Score 24; DB 13; Length 390;
Best Local Similarity 100.0%; Pred. No. 0.34;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 153 AAGAAAAAACAAAAACAAACACA 176
Db 75 AAGAAAAAACAAAAACAAACACA 98

RESULT 12

US-10-437-963-86792
; Sequence 86792, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 86792
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(467)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_857C.1
US-10-437-963-86792

Query Match 1.4%; Score 24; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 TTATATTATGGACGGAGGAAGTA 139
Db 373 TTATATTATGGACGGAGGAAGTA 396

RESULT 13

US-10-437-963-43365/c
; Sequence 43365, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43365
; LENGTH: 1480
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46529C.1
US-10-437-963-43365

RESULT 15
 US-10-437-963-4401/c
 ; Sequence 4401, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 4401
 ; LENGTH: 433

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:05:33 ; Search time 11065 Seconds
(without alignments)
6941.153 Million cell updates/sec

Title: US-10-009-570-1
Perfect score: 1772
Sequence: 1 acttagataataaagtaag.....ctcttcctccagtgcaag 1772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1772	100.0	1772	6	AX063520	Sequence
2	1772	100.0	1803	6	AX063522	Sequence
3	1686	95.1	1686	6	AX063525	Sequence
4	947	53.4	168173	8	AP004988	Oryza sat
5	660	37.2	2000	6	AX656635	Sequence
6	86	4.9	86	6	AX063523	Sequence
7	55	3.1	2771	8	AK100306	Sequence
8	52	2.9	139653	2	AP005890	Oryza sat
9	52	2.9	163555	2	AP005912	Oryza sat
10	50	2.8	135594	8	AC087220	Oryza sat
11	48	2.7	124366	8	AC131374	Oryza sat
12	48	2.7	150465	8	AC091749	Oryza sat
13	48	2.7	300029	8	AE017076	Oryza sat
14	45	2.5	2000	6	AX656219	Sequence
15	45	2.5	4817	8	AF254558	Oryza sat
16	45	2.5	139043	2	AC090974	Oryza sat
17	45	2.5	183580	8	AP003561	Oryza sat
18	44	2.5	137852	8	AP003880	Oryza sat
19	40	2.3	127506	8	OSJN00096	Oryza sat
20	40	2.3	167269	8	OSJN00085	Oryza sat
21	38	2.1	112483	2	AP004060	Oryza sat
22	38	2.1	132151	8	AC144491	Oryza sat
23	38	2.1	149807	2	AC120533	Oryza sat
24	38	2.1	151517	2	AP005750	Oryza sat
25	38	2.1	158291	8	AP004776	Oryza sat
26	38	2.1	159290	2	AP004813	Oryza sat
27	38	2.1	160562	2	AP005613	Oryza sat
28	38	2.1	168430	8	AC126221	Oryza sat
29	37	2.1	158815	8	AP003105	Oryza sat
30	37	2.1	164593	8	AP003197	Oryza sat
31	36	2.0	160652	2	AC134933	Oryza sat
32	36	2.0	135670	8	AP005002	Oryza sat
33	36	2.0	152058	8	AP005740	Oryza sat
34	36	2.0	175447	2	AC136956	Oryza sat
35	35	2.0	180544	2	AC137623	Oryza sat
36	35	2.0	121388	8	AP003759	Oryza sat
37	35	2.0	127506	8	OSJN00096	Oryza sat
38	35	2.0	133800	8	AP004801	Oryza sat
39	35	2.0	139482	2	AC113333	Oryza sat
40	35	2.0	149646	2	AP003761	Oryza sat
41	35	2.0	159626	8	AP005803	Oryza sat
42	35	2.0	173522	8	OSJN00060	Oryza sat
43	34	1.9	2000	6	AX656261	Sequence
44	34	1.9	131184	8	AP003868	Oryza sat
45	34	1.9	135014	8	AP003222	Oryza sat

ALIGNMENTS

RESULT 1
AX063520
LOCUS AX063520 1772 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 1 from Patent WO0078975.
ACCESSION AX063520
VERSION AX063520.1 GI:12541259
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Donaldson, I.A. and Rasmussen, T.B.
TITLE Rice sucrose synthase promoter
JOURNAL Patent: WO 0078975-A 1 28-DEC-2000;
DANISCO A/S (DK)

AUTHORS	Donaldson, I.A. and Rasmussen, T.B.
TITLE	Rice sucrose synthase promoter
JOURNAL	Patent: WO 0078975-A 6 28-DEC-2000; DANISCO A/S (DK)
FEATURES	Location/Qualifiers 1..1686 /organism="synthetic const /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="promoter sequence"

ORIGIN

Query Match	95.1%	Score 1686	DB 6	Length 1686
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1686	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy 1	ACITTTAGATATAAGTAAGTACACAGAGAAAATAATAATTCGAAATTTTTTTTATA 60			
Db				
Qy 1	ACITTTAGATATAAGTAAGTACACAGAGAAAATAATAATTCGAAATTTTTTTTATA 60			
Db				
Qy 61	AGACGAGTGGTCAAAACAGTCAAGTAAAACTCAAAATTCCTATATTATGGAGCTTATA 120			
Db				
Qy 61	AGACGAGTGGTCAAAACAGTCAAGTAAAACTCAAAATTCCTATATTATGGAGCTTATA 120			
Db				
Qy 121	TTATGGACGGAGGAAGTAGAAGATTGTGACCAAGAAAAACAAAAACAAACACACCGC 180			
Db				
Qy 121	TTATGGACGGAGGAAGTAGAAGATTGTGACCAAGAAAAACAAAAACAAACACACCGC 180			
Db				
Qy 181	CACCTGCGAGGCATGCATCTTAGTCCGCACATTTGAGAGGTCCGCAGTAGACGAGTTACC 240			
Db				
Qy 181	CACCTGCGAGGCATGCATCTTAGTCCGCACATTTGAGAGGTCCGCAGTAGACGAGTTACC 240			
Db				
Qy 241	CTACACAACTGCTTCTTCAGTGAGCTAGCTGCATGTTCTGTCTTCGCATTTACATTTGACGG 300			
Db				
Qy 241	CTACACAACTGCTTCTTCAGTGAGCTAGCTGCATGTTCTGTCTTCGCATTTACATTTGACGG 300			
Db				
Qy 301	CAGCAGCTAGCAACAGTTTGCAGGAACAATCGATAATCCATTGTGTCAAGGAGGAACAATG 360			
Db				
Qy 301	CAGCAGCTAGCAACAGTTTGCAGGAACAATCGATAATCCATTGTGTCAAGGAGGAACAATG 360			
Db				
Qy 361	GAGAAAAACCGGGCTGGAGACGACGGAGCAGCTGTACCGTAGCTTCTGAAAGGCTGA 420			
Db				
Qy 361	GAGAAAAACCGGGCTGGAGACGACGGAGCAGCTGTACCGTAGCTTCTGAAAGGCTGA 420			
Db				
Qy 421	ACCCATCTGCGAAATCGCAGATTGTTGTTTCAATTCCAACTGTCAGTCCCTTCAGATTG 480			
Db				
Qy 421	ACCCATCTGCGAAATCGCAGATTGTTGTTTCAATTCCAACTGTCAGTCCCTTCAGATTG 480			
Db				
Qy 481	GTTCATGTTTCAACCGTAGTACATCTGAAAAATGAAGTGTAAATPACCTTGAGAGACCT 540			
Db				
Qy 481	GTTCATGTTTCAACCGTAGTACATCTGAAAAATGAAGTGTAAATPACCTTGAGAGACCT 540			
Db				
Qy 541	TCATGGAAGCATGCTGCAGCGGATTAGCTAAGAAAAAATAAATGTACTTTTCGAA 600			
Db				
Qy 541	TCATGGAAGCATGCTGCAGCGGATTAGCTAAGAAAAAATAAATGTACTTTTCGAA 600			
Db				
Qy 601	ACTTAAATTTGAGTTAGATTTTAGGGTGTTTCCATCGTAGTGATTTTCTACTATTGCA 660			
Db				
Qy 601	ACTTAAATTTGAGTTAGATTTTAGGGTGTTTCCATCGTAGTGATTTTCTACTATTGCA 660			
Db				
Qy 661	GTTTAAACCGCTAATAGTCAGATATAAAATTTTATCTATAGATCAATTTATAAATCATTTT 720			
Db				
Qy 661	GTTTAAACCGCTAATAGTCAGATATAAAATTTTATCTATAGATCAATTTATAAATCATTTT 720			
Db				
Qy 721	TAGTTGCTTCGTTTCATTTTCTACACTTATCAACCATAGCTCAACTGATCAATTTGACAA 780			
Db				
Qy 721	TAGTTGCTTCGTTTCATTTTCTACACTTATCAACCATAGCTCAACTGATCAATTTGACAA 780			
Db				
Qy 781	TAAAGTTTACTTAAACGACATCGCTCATCACACCCACCGCTCACCGATGGTGCTCTC 840			
Db				
Qy 781	TAAAGTTTACTTAAACGACATCGCTCATCACACCCACCGCTCACCGATGGTGCTCTC 840			
Db				
Qy 841	GACCACGAGTTTAGCACCTTGTGCAACATATATGCGTGGATGAAATCTACTGATGCGCG 900			
Db				

RESULT 4	ACCESSION
AP004988	VERSION
LOCUS	KEYWORDS
DEFINITION	SOURCE
	ORGANISM

	Db	841	GACCAAGATTTAGCACTTGTGCAACAATATATCGTGCGATGAACATCTACTGATGGCC	900
	Qy	901	ATCGGAATTTTAGCGTTTCGTTTCATGACGCTTCCAAACGGCACAAGAGCTGAGCAGCAGCAT	960
	Db	901	ATCGGAATTTTAGCGTTTCGTTTCATGACGCTTCCAAACGGCACAAGAGCTGAGCAGCAGCAT	960
	Qy	961	GCATGCATGGCTCTCTGTAACACAAAAGAAGTTTACTTGGTTAAATGACATGCTGCTGTAGCT	1020
	Db	961	GCATGCATGGCTCTCTGTAACACAAAAGAAGTTTACTTGGTTAAATGACATGCTGCTGTAGCT	1020
	Qy	1021	AGTTAGCAGAAATGCAAGGCCCATGCAATATGCAATGCTATGCAACAAGTATAGTACCAGCA	1080
	Db	1021	AGTTAGCAGAAATGCAAGGCCCATGCAATATGCAATGCTATGCAACAAGTATAGTACCAGCA	1080
	Qy	1081	TGTATGGTAGCCAGCTTAACCTAATCTATCAGCAGAGCAGCAAGCTCGTGATGGTGTGAT	1140
	Db	1081	TGTATGGTAGCCAGCTTAACCTAATCTATCAGCAGAGCAGCAAGCTCGTGATGGTGTGAT	1140
	Qy	1141	GCATCTCTCCAGTAATCTTAGTGGTAAATTTTTCACCCAAGCGTTGCTCATATGGACAGT	1200
	Db	1141	GCATCTCTCCAGTAATCTTAGTGGTAAATTTTTCACCCAAGCGTTGCTCATATGGACAGT	1200
	Qy	1201	AATTAGTAATATTACCAAGGTTACAAATCCGTTTACTGACCAAAATACTACTCACGAATG	1260
	Db	1201	AATTAGTAATATTACCAAGGTTACAAATCCGTTTACTGACCAAAATACTACTCACGAATG	1260
	Qy	1261	GTATCTCTGGTTTTCTGTTAAAAACCGTTGGTTAAACCCAGCAAAAATAGACAAAATTTGTCAA	1320
	Db	1261	GTATCTCTGGTTTTCTGTTAAAAACCGTTGGTTAAACCCAGCAAAAATAGACAAAATTTGTCAA	1320
	Qy	1321	AATTTTAAATTTTAGTTTTTTTTTTTAACTTAGCGGGGAAACCTTGAAGTTTGTGCTGTGCG	1380
	Db	1321	AATTTTAAATTTTAGTTTTTTTTTTTAACTTAGCGGGGAAACCTTGAAGTTTGTGCTGTGCG	1380
	Qy	1381	AGCTGTCTCGGAAGACGGTTTTTGGTTGGGATTTGTGAACCCCTGGTTACTGCACTTCAAT	1440
	Db	1381	AGCTGTCTCGGAAGACGGTTTTTGGTTGGGATTTGTGAACCCCTGGTTACTGCACTTCAAT	1440
	Qy	1441	TTTGAACAGATATTAGTGCAACAGACAAATGCAACGCAATTTTTTCTGTTTACCGGCAA	1500
	Db	1441	TTTGAACAGATATTAGTGCAACAGACAAATGCAACGCAATTTTTTCTGTTTACCGGCAA	1500
	Qy	1501	GCTGMAAGCTTTTAGATCCCCATACCGCGCTGCTGCAAAACCTGCCAAGAAAGACGACGA	1560
	Db	1501	GCTGMAAGCTTTTAGATCCCCATACCGCGCTGCTGCAAAACCTGCCAAGAAAGACGACGA	1560
	Qy	1561	GAACACAGGTGTCAATTTTGTGGTGGTAAAGCCCAAGTAAAGTAAACAGAAAGTGGAGATAGT	1620
	Db	1561	GAACACAGGTGTCAATTTTGTGGTGGTAAAGCCCAAGTAAAGTAAACAGAAAGTGGAGATAGT	1620
	Qy	1621	GAGGACACAGGAGTAGGCGAGGGGACATATGGCCCCACGCTCCCTGACCAATTTTCGTGTA	1680
	Db	1621	GAGGACACAGGAGTAGGCGAGGGGACATATGGCCCCACGCTCCCTGACCAATTTTCGTGTA	1680
	Qy	1681	TAAATA 1686	
	Db	1681	TAAATA 1686	

AP004988 168173 bp DNA linear PLN 23-JUL-2003
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
BAC clone: B1056G08.
AP004988
AP004988.3 GI:24060020
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

1
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2
Published Only in Database (2002)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
clone.B1056G08
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (27-MAR-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasa@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Oct 16, 2002 this sequence version replaced gi:22212589.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM
(http://www.tigr.org/tdb/glimmerm/glmr form.html), RiceHMM
(http://rgp.dna.affrc.go.jp/RiceHMM/) -SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/hmli/docs/sim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBJ. Protein homologies of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This is a partial sequence of B1056G08 clone. This sequence
of B1056G08 clone has an overlap with P0616D06 (DBJ: AP005198) at
5' end and an overlap with P0552F09 clone (DBJ: AP004309) at 3' end
and an overlap with P0560B08 clone (DBJ: AP004309) at 3' end.
Detailed information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.

Location/Qualifiers
1. 168173
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="B1056G08"
complement(3083..4070)
/gene="B1056G08.101"
complement(join(3083..3490,3648..4070))
/gene="B1056G08.101"
/note="supported by full-length cDNA(s): AK108210"
complement(join(3286..3490,3648..4036))
/gene="B1056G08.101"
/note="contains Set's): AU069911(E12068)
contains full-length cDNA(s): AK108210"
/codon_start=1
/product="unknown protein"
/protein_id="BAC21478.1"
/db_xref="GI:24060021"
/translation="MASAVVAKPPAAPASRRRCFFVVERAAAAAVGVNAAIAVDDH
RPVPAAAAAAMDDVUGRVARPARPSARAIWEGTHKQISSGGASGICYTVPWCSTICN

NPFAIAEFLLCNLGVPLAGRPSFIYIGKAFCKBECRSRYVVVEALREAREEKRA
AAAAASPEKKKAAARKGEGREGSIFICADLL"
complement(join(5482..5716,6013..6023))
/gene="B1056G08.102"
complement(join(5482..5716,6013..6023))
/gene="B1056G08.102"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
join(7413..7432,7688..8096)
/gene="B1056G08.103"
join(7413..7432,7688..8096)
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
8222..8962
/gene="B1056G08.104"
join(<8222..8368,8759..>8962)
/gene="B1056G08.104"
/note="start and end point are not identified"
join(8222..8368,8759..8962)
/gene="B1056G08.104"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC21479.1"
/db_xref="GI:24060022"
/translation="MDLSAGRRASFPSPPLAAAIAPSPSSPADRAPLPRRIT
APPDNNWTKVTGSRAPNPNFGHIKLRQSLDLGATSVTQFTNYITCRSSPSDNT
GYEHKNGQDSAP"
complement(8972..11317)
/gene="B1056G08.105"
complement(join(<8972..9052,9394..9552,11110..11185,
11283..>11317))
/gene="B1056G08.105"
/note="start and end point are not identified"
complement(join(8972..9052,9394..9552,11110..11185,
11283..11317))
/gene="B1056G08.105"
/note="contains Set's): C74017(E30162)"
/codon_start=1
/product="putative RNA Polymerase II subunit 14.5 kD"
/protein_id="BAC21480.1"
/db_xref="GI:24060023"
/translation="MAMSAKFCGECNMILYPREDKETHLLYACNSCEHQEPATDFC
VYKRVLRKPAGEPKDILKDAATDPTLPTRSIKYNGCHPEAAFFOAPTKEGKLTLY
FICCNPSCGHRWRD"
join(11555..11934,12176..12278,12484..14337,14423..14478,
14801..15101)
/gene="B1056G08.106-1"
join(<11555..11934,12176..12278,12484..14246,14423..14478,
14801..>15101)
/gene="B1056G08.106-1"
/note="supported by full-length cDNA(s): AK073803"
join(11585..11934,12176..12278,12484..14337,14423..14478,
14801..15636)
/gene="B1056G08.106-2"
join(<11585..11934,12176..12278,12484..14337,14423..14478,
14801..>15636)
/note="supported by full-length cDNA(s): AK066317"
join(11828..11934,12176..12278,12484..14246,14423..14478,
14801..14916)
/gene="B1056G08.106-2"
/note="putative transposase"
join(11828..11934,12176..12278,12484..14337,14423..14473)
/gene="B1056G08.106-1"
/note="putative transposase"
21538..28635
/gene="B1056G08.107"
join(<21538..21596,21679..21859,21962..22064,22185..22240,
23693..23845,24429..26511,26756..28073,28346..28429,

FEATURES
source

gene

mRNA

CDS

28571. >28635)	Db	67804	AACCGGGCTGGAGACGAAACGGAGCAGCTGTACCGTACGTTCTTGAAGGCTGAACCCAT	67863
/gene="B1056G08.107"				
/note="start and end point are not identified"				
Join(21538. .21596,21679. .21858,21962. .22064,22185. .22240,23693. .23845,24429. .26511,26756. .28073,28346. .28429,28571. .28635)	Qy	427	CTGCGAATTCGCGAGATTGGTTTCAATTCACACTTGCAGTCCCTTCAGATTGGTTGCA	486
/gene="B1056G08.107"	Db	67864	CTGCGAATTCGCGAGATTGGTTTCAATTCACACTTGCAGTCCCTTCAGATTGGTTGCA	67923
/note="contains Set's": C26549 (C12568), AU092341 (C12568)	Qy	487	TGTT--CAACCCGTAGTACATCTCTGAAATGAAGTGTAAATACCTTTGAGAAGACCTTCATG	545
contains full-length cDNA(s): AK109921"				
/codon_start=1	Db	67924	TGTTCCACCGTAGTACATCTGAAATGAAGTGTAAATACCTTTGAGAAGACCTTCATG	67983
/product="COPI-interacting protein 7 (CIP7) -like"	Qy	546	GAAGCATGCTGCGAGCGGATAGCTAAGAAAAAATAATGATGACTTTTTCGAAACTTA	605
/protein_id="BAC21482.1"				
/db_xref="GI:24060025"	Db	67984	GAAGCATGCTGCGAGCGGATAGCTAAGAAAAAATAATGATGACTTTTTCGAAACTTA	68043
/translation="MEPADLPALFQSPRRSRCELVSNGRTERIASGVKPFVA	Qy	606	ATTTTGGAGTTAGATTTTAGGGTGTTCATCGTAGTGTATTTTCTACTATTGCGAGTTTA	665
HLRAEEAQAQPPPPAQLQLDREAAMFSGTGLRFFVSTPEVLAMANTFDEAMS	Db	68044	ATTTTGGAGTTAGATTTTAGGGTGTTCATCGTAGTGTATTTTCTACTATTGCGAGTTTA	68103
QLEGARKIAQVAGGAGAGAHUNVYVLLKELLRAIDLRLSALKQDLVTCARASS	Qy	666	AACCGCTAATAGTCAGATATAAAATTTTATCTATAGATCAATTTATAAATCAATTTAGTT	725
AGFPDVSVELFADHFGAKSEACFMWLCORREDI CHYISVSTSGCWKSFDD	Db	68104	AACCGCTAATAGTCAGATATAAAATTTTATCTATAGATCAATTTATAAATCAATTTAGTT	68163
GNVSSSSDLSDTDQAGASSNKSIIIGSVSHRSNSQNSVDPPEPSAVQHPK	Qy	726	GCTTCGTTCAATTTTCTACCACTTATCAACATAGCTCAACCGATCAATTTGAACAATAAA	785
PTIQQSVQKEKETDALPAPAGGGSRLSVQDRINMFESKQKEQTS SGNAACTS	Db	68164	GCTTCGTTCAATTTTCTACCACTTATCAACATAGCTCAACCGATCAATTTGAACAATAAA	68223
KVPTKGHRVPVSGMDKLVRWSNVDSMISLNSNDSSLNKEIKETIGTVPVPTS	Qy	786	GTTTACTAAACGACATCGCTCATCACACACCCACCTCACCGATGGTGGCTCTCGACCA	845
ANLEVSKARADGDKLVKHTVSCQKDTSDALPLDSTADAFSSSTLNTTSPSLA	Db	68224	GTTTACTAAACGACATCGCTCATCACACACCCACCTCACCGATGGTGGCTCTCGACCA	845
IASSPQKTPRVEDDMVITSSIESESFRKEVASOGKGVNRSQGVASVSTRAR	Qy	846	CGAGTTTAGCACTTGTGCAACATATATGCGTGCATGAACATCTACTGATGCCATGCG	905
VKTSPRTWPNNTVLSPPULSGEHVQMTDEETIPIVHEAVVKEQI VOKNRRGLRL	Db	68284	CGAGTTTAGCACTTGTGCAACATATATGCGTGCATGAACATCTACTGATGCCATGCG	905
SKETHAEADVVRKDRPRTTKIISDTRFRATSNPRNFRSGSVDEASTAEVHVD	Qy	906	AATTTTAGCGTTCGTTCTATGACGCTTCCAAACGCGACAGAGGCTGAGCAGCAGCATG	965
NLQKSLARKVEDSKVAAGSEILPQSDCTSHOCTNLSRQSSSAEQLSHGGVKL	Db	68344	AATTTTAGCGTTCGTTCTATGACGCTTCCAAACGCGACAGAGGCTGAGCAGCAGCATG	965
ISDNVAPLEOTKPTKGSQDRHDLQKANELEKFAAKLTSSRRKSTDVQVNT	Qy	966	CATGCTCTCTGTAAGAAACAAAAGAGTTTACTGCTGTAATGACATGCTGCTGTAGTAGTTA	1025
PRNVKPLVLTPIRYTTOIVKESI TNEFANEILLKMDTGYNNYPOSTIISLEES	Db	68404	CATGCTCTCTGTAAGAAACAAAAGAGTTTACTGCTGTAATGACATGCTGCTGTAGTAGTTA	1025
RKGFYDMQKRDALKEDWKLQKQKEATIKMRDLSLERSNAEMRAKESRSSVPS	Qy	1026	GCAGATGCAAGGCCCATGCATATGCAATGCTATGCAACAGTATAGTACCAAGCATGTAT	1085
TIYSIRCAHFPPLQVQKQIDISFLVEEMNSDYLSGGDSRSASDRKHFKNKVA	Db	68464	GCAGATGCAAGGCCCATGCATATGCAATGCTATGCAACAGTATAGTACCAAGCATGTAT	1085
CNQKSLAPVHRSSRTVSSGYANRNLPDNLQSVNFDLRKENTKPSAGLSRAA	Qy	1086	GGTAGCCAGCTAACTTAATCTATCAGCAGAGGCGACAGCTCGTGCATGGTGTGATGCAT	1145
PTOPKPSIRKSIITEESKNISKQSRKNSMRKNLSPELADTSMNDIYNWAPSK	Db	68524	GGTAGCCAGCTAACTTAATCTATCAGCAGAGGCGACAGCTCGTGCATGGTGTGATGCAT	1145
ISNDVEGVFAVITHTAGSTKSLFKAGNEHPAVGAGPAPMPFANTYQNGDDDFLD	Qy	1146	TCCTCCAGTAACTTAGTGGTAAATTTTCAACCAAGCGTTGCTCATATGACAGTAATTA	1205
OEEDSPDKBEYESIENLRESDFPADSDENPGISHEFGNSDDPGSENGDVSPFS	Db	68584	TCCTCCAGTAACTTAGTGGTAAATTTTCAACCAAGCGTTGCTCATATGACAGTAATTA	1205
DAPTLGCKFNAPAGNMDTPEGVPAASHSTRPLPAYANDNSDGFADSPNGSPSP	Qy	1206	GTAATATTACCAAGGTTCAAAATCCCGTTACCTGACCAATACTACTCTACGAATGTATC	1265
NSHTDQITDADVRMRKKGWQAPFVGNPASQPRKQDITKGFKLKFGKRTKAD	Db	68644	GTAATATTACCAAGGTTCAAAATCCCGTTACCTGACCAATACTACTCTACGAATGTATC	1265
GLNDVVSATASECDDHEDGLDLAGSSDDFRKMRGYPGSDYDGFVDTFVAFEDQS	Qy	1266	TCGTGGTTTCGTTTAAACCGTTGGTAAACCGTAAACAGCAAAAATAGCAAAAATTTTGC	1325
LRSSIPNPNPFRLEDOLITSSLKAPRSFFSLSTRSKGGDARLR	Db	68704	TCGTGGTTTCGTTTAAACCGTTGGTAAACCGTAAACAGCAAAAATAGCAAAAATTTTGC	1325
29193. .31064	Qy	1326	TAAATTTTAG--TTTTTTTTTTAACTTAGCCGGGAAACCTTCAAGTGTGCTGTCGACG	1383
/gene="B1056G08.108"				
<29193. .>31064	Db	68764	TAAATTTTAGTTTATTTTATTTTAACTTAGCCGGGAAACCTTCAAGTGTGCTGTCGACG	68823
/note="supported by full-length cDNA(s): AK064501"	Qy	1384	TGCTCTGGGAAGACGTTTGGTGGGATTTGGATTTGGAAACCTGCTGCTACTGCACTTTT	1443
29193. .31064	Db	68824	TGCTCTGGGAAGACGTTTGGTGGGATTTGGTGGGATTTGGAAACCTGCTGCTACTGCACTTT	68883
misc_feature	Qy	1444	GAACAGATATTAGTCACAGCAACAAATGCAACGATTTTTTCTGTTTTACCGGCAAGCT	1503
Query Match 53.4%; Score 947; DB 8; Length 168173;				
Best Local Similarity 99.3%; Pred. No. 0;				
Matches 1747; Conservative 0; Mismatches 10; Indels 3; Gaps 2;				
7 GATAATAAGTAGTACGACAGAAATAATAATAATTCCTCAAAATTTTTTAAAGACGA 66	Qy			
	Db	67444	GATAATAAGTAGTACGACAGAAATAATAATAATTCCTCAAAATTTTTTAAAGACGA 67503	
67 GTGTGCAACAGTACAGTAAACCTCAAAATTCCTTATATATGCGACTTATATATGCG 126	Qy			
	Db	67504	GTGTGCAACAGTACAGTAAACCTCAAAATTCCTTATATATGCGACTTATATATGCG 67563	
127 GACGAGGAGTAGAAGATTGTAGCCACGAAAAACAAAAACAAACACACCCGACCTG 186	Qy			
	Db	67564	GATGAGGAGTAGAAGATTGTAGCCACGAAAAACAAAAACAAACACACCCGACCTG 67623	
187 GCAGGCGATGCATCTTAGTTCGGCACAATTGAGAGGTCGGCAGTAGCAGTTACCCCTACAC 246	Qy			
	Db	67624	GCAGGCGATGCATCTTAGTTCGGCACAATTGAGAGGTCGGCAGTAGCAGTTACCCCTACAC 67683	
247 AACTGCTCTTCAGTGAGTACGATGCTTCTGTTCTGCATTTACATTTGCGAGGACGAG 306	Qy			
	Db	67684	AGCTGCTCTTCAGTGAGTACGATGCTTCTGTTCTGCATTTGCAATTTGCGAGGACGAG 67743	
307 CTAGCAACAGTTTTCAGGAAACCAATCGATAATCAATTTGTTGTCAGGAGGAAACATGGAGAAA 366	Qy			
	Db	67744	CTAGCAACAGTTTTCAGGAAACCAATCGATAATCAATTTGTTGTCAGGAGGAAACATGGAGAAA 67803	
367 AACCGGGCTGGAGACGAAACGGGAGCAGCTGTACCGTACGTTTCTGAAAGCTGAACCCAT 426	Qy			

Db 68884 GAACAGATATTAGTGCAACAGACAAATGTCACCAAGCAATTTTTTCTGTTTACCGGCAAGCT 68943
Qy 1504 GAAGCTTTTACGATCCCATACCGCGCTTGTCTGCAAAACCTGCCAAGAAAGAGCAGAGAA 1563
Db 68944 GAAGCTTTTACGATCCCATACAGCCGTTGCTGCAAAACCTGCCAAGAAAGAGCAGAGAA 69003
Qy 1564 ACAGCTGTCAATTTTGTGGTGAAGAGCCAAAGTAAAGTAAACAGAAAGATGGAAGATAGTGAG 1623
Db 69004 ACAGCTGTCAATTTTGTGGTGAAGAGCCAAAGTAAAGTAAACAGAAAGATGGAAGATAGTGAG 69063
Qy 1624 GACCAGGAGTGAAGAGGAGGAGACACATGGGCCACCGCTCCCTGCACATTTTCGTGTATAA 1683
Db 69064 GACCAGGAGTGAAGAGGAGGAGACACATGGGCCACCGCTCCCTGCACATTTTCGTGTATAA 69123
Qy 1684 ATACAGTGTGATGCATCGCTCTCCAGCATCCATCGGTTCTCTGCTCTGTTTCATCCATAG 1743
Db 69124 ATACAGTGTGATGCATCGCTCTCCAGCATCCATCGGTTCTCTGCTCTGTTTCATCCATAG 69183
Qy 1744 AGTTTCCTCTCTCTCTCTCTT 1763
Db 69184 AGTTTCCTCTCTCTCTCTT 69203

RESULT 5
AX656635
LOCUS 2000 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 6505 from Patent WO03000898.
ACCESSION AX656635
VERSION AX656635.1 GI:29159449
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Qian, S., Qian, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 6505 03-JAN-2003;
Syngenta Participations AG (CH)
Location/Qualifiers
1. .2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Query Match 37.2%; Score 660; DB 6; Length 2000;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1090; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
Qy 669 CGCTAATAGTCAGATATAAAATTTATCTATAGATCAATTTATAAATCAATTTTAGTTGCT 728
Db 1 CGCTAATAGTCAGATATAAAATTTATCTATAGATCAATTTATAAATCAATTTTAGTTGCT 60
Qy 729 TCCTTCATTTTCTACCACTATCAACCATAGCTCAATCTGATCAATTCAGAAATGATT 788
Db 61 TCCTTCATTTTCTACCACTATCAACCATAGCTCAACCGATCAATTCAGAAATGATT 120
Qy 789 ACTAAACGACATCGCTCATCACACACCAACCGCTCAGCGATGGTGGCTCTCGACACACGA 848
Db 121 ACTAAACGACATCGCTCATCACACACCAACCGCTCAGCGATGGTGGCTCTCGACACACGA 180
Qy 849 GTTTAGCACTTGTGCAACATATATGCGTGGGATGAACATCTACTGATGCGCCATGCGAAT 908
Db 181 GTTTAGCACTTGTGCAACATATATGCGTGGGATGAACATCTACTGATGCGCCATGCGAAT 240
Qy 909 TTTTAGCGTTCATGACGCTTCCACGCGACAGAGCTGAGCAGCAGATGATGATCAT 968
Db 241 TTTTAGCGTTCATGACGCTTCCACGCGACAGAGCTGAGCAGCAGATGATGATCAT 300
Qy 969 GGCTCTTGTGAAACAAAGGTTACTGGTAAATGACATGCTGCTAGCTAGTTAGCA 1028

Db 301 GGCTCTTGTGAAACAAAGGTTACTGGTAAATGACATGCTGCTGTAGCTAGTAGCA 360
Qy 1029 GAATGCAAGGCCCATGATATGCAATGCTATGCAACAGTATAGTACAGCATGTATGGT 1088
Db 361 GAATGCAAGGCCCATGATATGCAATGCTATGCAACAGTATAGTACAGCATGTATGGT 420
Qy 1089 AGCCAGTAACTAATCTATCAGCAGAGGAGCAGCTCGTGCATGGTGTGATGCACTTCT 1148
Db 421 AGCCAGTAACTAATCTATCAGCAGAGGAGCAGCTCGTGCATGGTGTGATGCACTTCT 480
Qy 1149 CTCAGTAACTAGTGTAAATTTTCAACCAAGGTTGCTCATATGACAGTAAATAGTA 1208
Db 481 CTCAGTAACTAGTGTAAATTTTCAACCAAGGTTGCTCATATGACAGTAAATAGTA 540
Qy 1209 ATATTCAAGGTTCAATCCCGTTTACCTGACCAAAATCTACTCAGAAATGGTATCTCT 1268
Db 541 ATATTCAAGGTTCAATCCCGTTTACCTGACCAAAATCTACTCAGAAATGGTATCTCT 600
Qy 1269 GGTTCCTGTTAAACCGTTGGTAAACAGCAAAATAGCAAAATTTGTCAAAATTTTAA 1328
Db 601 GGTTCCTGTTAAACCGTTGGTAAACAGCAAAATAGCAAAATTTGTCAAAATTTTAA 660
Qy 1329 ATTTTAG- -TTTTTTTTTAACTTAGCCGGGAAACCTTGAAGTTTGTGCTGTGAGCTGT 1386
Db 661 ATTTTAGTTTTTTTTTAACTTAGCCGGGAAACCTTGAAGTTTGTGCTGTGAGCTGT 720
Qy 1387 CTTGGGAAAGCAGCGTTTGGTGGGATTGTGAACCCCTGGTTACTGCACTTCATTTTGA 1446
Db 721 CTTGGGAAAGCAGCGTTTGGTGGGATTGTGAACCCCTGGTTACTGCACTTCATTTTGA 780
Qy 1447 CAGATATTAGTCAACAGACAGAAATGCCAAGCATTTTTTCTGTTTACCGCAGCTGAA 1506
Db 781 CAGATATTAGTCAACAGACAGAAATGCCAAGCATTTTTTCTGTTTACCGCAGCTGAA 840
Qy 1507 GCTTTTACGATCCCATACCGCGTGTCTGCAAAACCTGCCAAGAGAGCAGCAAGAAACA 1566
Db 841 GCTTTTACGATCCCATACGCGTGTCTGCAAAACCTGCCAAGAGAGCAGCAAGAAACA 900
Qy 1567 GGTGTCAATTTTGTGGTGAAGCCAAAGTAAAGTAAACAGAAAGATGGAAGATAGTAGGAC 1626
Db 901 GGTGTCAATTTTGTGGTGAAGCCAAAGTAAAGTAAACAGAAAGATGGAAGATAGTAGGAC 960
Qy 1627 CAGGAGTGAAGGAGGAGACACATGGCCACGCTCCCTGCACATTTTCGTGTATAAATA 1686
Db 961 CAGGAGTGAAGGAGGAGACACATGGCCACGCTCCCTGCACATTTTCGTGTATAAATA 1020
Qy 1687 CAGTGTGATGCATCGCTCTCCAGCATCCATCGTTCTCTGCTCTGTTTCATCCATAGT 1746
Db 1021 CAGTGTGATGCATCGCTCTCCAGCATCCATCGTTCTCTGCTCTGTTTCATCCATAGT 1080
Qy 1747 TTCCTCTCTCTCTCTT 1763
Db 1081 TTCCTCTCTCTCTCTT 1097

RESULT 6
AX063523
LOCUS 86 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 4 from Patent WO0078975.
ACCESSION AX063523
VERSION AX063523.1 GI:12541262
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1 Donaldson, I.A. and Rasmussen, T.B.
Rice sucrose synthase promoter
Patent: WO 0078975-A 4 28-DEC-2000;
DANISCO A/S (DK)
Location/Qualifiers
1. .86
source

JOURNAL
REFERENCE 2 (bases 1 to 139653)
AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsaaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
Location/Qualifiers
1..139653
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="B1029C06"

ORIGIN
Query Match 2.9%; Score 52; DB 2; Length 139653;
Best Local Similarity 100.0%; Pred. No. 3.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 28 AAAATAAATAAATTCCTCAAAATTTTAAATAGACGAGTGGTCAACAGT 79
|||||
Db 27521 AAAATAAATAAATTCCTCAAAATTTTAAATAGACGAGTGGTCAACAGT 27572
|||||

RESULT 9
AP005912/c 163555 bp DNA linear HTG 26-NOV-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 9 clone
DEFINITION OSJNBa0064123, *** SEQUENCING IN PROGRESS ***.
ACCESSION AP005912.1 GI:253396710
VERSION HTG; HTGS PHASE2
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
clone:OSJNBa0064123
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 163555)
AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsaaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES
source
Location/Qualifiers
1..163555
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="OSJNBa0064123"

ORIGIN
Query Match 2.9%; Score 52; DB 2; Length 163555;
Best Local Similarity 100.0%; Pred. No. 3.2e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 28 AAAATAAATAAATTCCTCAAAATTTTAAATAGACGAGTGGTCAACAGT 79
|||||
Db 55781 AAAATAAATAAATTCCTCAAAATTTTAAATAGACGAGTGGTCAACAGT 55730
|||||

RESULT 10
AC087220/c 135594 bp DNA linear PLN 12-MAR-2003
LOCUS Oryza sativa chromosome 3 BAC OSJNBb0097F01 genomic sequence,
DEFINITION complete sequence.
ACCESSION AC087220
VERSION AC087220.9 GI:28460790
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 135594)
AUTHORS Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, K.M.,
Overton II, L., Tsitrin, T., Kim, M.M., Bera, J.J., Jin, S.S.,
Fadrosh, D.W., Tallon, L.J., Koo, H., Zismann, V., Heiao, J., Blunt, S.,
Vanaken, S.S., Riedmuller, S.B., Utterback, T.T., Feldblum, T.V.,
Yang, Q.Q., Haas, B.J., Suh, B.B., Peterson, J.J., Quackenbush, J.,
White, O., Salzberg, S.L. and Fraser, C.M.
TITLE Oryza sativa chromosome 3 BAC OSJNBb0097F01 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 135594)
AUTHORS Buell, R.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 135594)
AUTHORS Buell, R.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 135594)
AUTHORS Buell, R.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuelletigr.org
On Feb 21, 2003 this sequence version replaced gi:16418178.
Address all correspondence to:rice@tigr.org
BAC clone OSJNBb0097F01 is from Oryza sativa chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (<http://www.softberry.com/>),
GENSCAN and GENSCAN+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkEM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and GeneSplicer
(Mihaela Pertea and Steven Salzberg, [contact mper@tigr.org](http://mper.tea.tigr.org/)),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named


```
gene
/rpt family="AT rich"
complement(20634..21896)
/gene="OSJNB0097F01.6"

Query Match      2.8%; Score 50; DB 8; Length 135594;
Best Local Similarity 100.0%; Pred. No. 4e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GATAATAAGTCAACAGAAAAATAATAATAATTCCAAATTTTTT 56
    |||
Db 52161 GATAATAAGTCAACAGAAAAATAATAATAATTCCAAATTTTTT 52112

RESULT 11
AC1131374 LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 10 clone
OSJNB0008A05, complete sequence.
AC1131374
AC1131374.1 GI:22380696
HTG.
ORIGIN Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 124366)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
Currie,J., and Collura,K.
Rice Genomic Sequence
Unpublished
2 (bases 1 to 124366)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
Currie,J., and Collura,K.
Direct Submission
Submitted (21-AUG-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
3 (bases 1 to 124366)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H., Rambo,T., Saski,C.,
Currie,J., Collura,K. and Thurmond,S.K.
Direct Submission
Submitted (03-SEP-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
The following sequence is an artificial clone that will be used for
gap filling in the rice genome. There is 100 kb of sequence from
the overlapping north clone (OSJNB0008A05) and 7982 bases of
sequence from the overlapping south clone (OSJNB0027A14) and the
gap filling sequence begins at 100,001 and ends at 116,384 for a
total size of 124,366 bases. The gap filling sequence is 16,384
bases. The name (OSJNB0008A05) was derived from the location of
the north clone and the 'A' stands for Artificial.
LOCATION/Qualifiers
1..124366
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/sub_species="(japonica cultivar-group)"
/db_xref="taxon:39947"
/chromosome="10"
/clone="OSJNB0008A05"
1994..2011
/note="Simple Sequence Repeat (TC)n"
3767..4477
/gene="OSJNB0008A05.1"
/note="Contains similarity to receptor-like protein
kinase"
join(3767..3924,4300..4477)
/gene="OSJNB0008A05.1"
/codon_start=1
/product="Putative receptor-like protein kinase"
/db_xref="GI:22655731"
/translation="MAAANAAGAEDSGELKGGSVWRKNGEALPLPEWEGNES
LAQVARLGWVLOGCRAGVQGNQAOQWHEHFQHEVETTRTWRGELVDALKLTQLFLIG
```

```
gene
complement(6525..14193)
/gene="OSJNB0008A05.2"
/note="Contains similarity to receptor-like protein
kinase"
CDS
7716..7926,7998..8122,9355..9443,9532..9772,11989..12084,
13110..13229,13335..14193)
/gene="OSJNB0008A05.2"
/codon_start=1
/product="Putative receptor-like protein kinase"
/protein_id="AA04149.1"
/db_xref="GI:22655732"
/translation="MTVPTMLNSTLPLVLLAPLSPHCRQAVLNYSCNNGSSVA
ENSTINVRXAVLATLSASTNATTFATAGSAGADTVMGALCRGDTDRVACSCV
AAYPATAFAGRGVRDVTYFYDRCIARFSYDFTARDNTDTEVLMVSKYQATNAGH
FDALVTRLAGALADWRAYNTRYAGVMASDGFPTTGYVMHNYGLVQCADPLAP
PACRACIALIVDMPOAFGRIGGOFNAVCNLYRSETSVFYDGPAPVRLVASCPELGP
NINGSTTLIIGNRRRPNAAATVVSVAAVIALLSVLSIYLWRKLOAKOCSPLAATA
OQLIVQAASKQIKQSTYSKQAMLAVDGDNALGEDVGEVGLDECAPVVEGA
VVAALSVARGDGDGTRSAVLRHLMVAPKRSASWSSTAKAMPRPPRPTAASP
HRIGPAAASPRRHLTPADTADQDVEAGSLFLDLATLRKATANFAEYKVLGGFGAV
YKFLRDGEIEIAVRKLDKASQGTIEQLNELLLVAKLRHNNLAKLGVCKIGBEKLLV
YELPNRSLDTFLDPEKRLQIWTETRIHGTARGLVYHEDSHIKIHLRLKASN
VLIDSNMPKISDFGLARLFDGNKTASVTSHVVTGLGYMAPEVAILSVKLDVYSF
GLVLEVVTCGRNTDVGAVEESNNLLSYVDHWKGTPLAIVDASLLGDRGPPESE
MLKIQGLLCCVSEENPADRPTMLHLVNLHDVDTATSPAFSKPAFTFVNGHTTGSS
NVAALSLNEVSISEFHR"
16903..16930
/note="Simple Sequence Repeat (ATAG)n"
complement(20967..23213)
/note="Similar to truncator transposon"
complement(23976..28321)
/gene="OSJNB0008A05.3"
/note="Contains similarity to polyprotein"
complement(join(23976..27756,27852..28321))
/gene="OSJNB0008A05.3"
/codon_start=1
/product="Putative polyprotein"
/protein_id="AA04150.1"
/db_xref="GI:22655733"
/translations="MVNTRSGNPNNNNNNGENTLAQVLAQQAQLMMMMQQLQ
NQOQNNHAPPONKLAELRVPPPIFSSNTNVEAGDWLHATEKLDLLQCTQENV
SPASHLHGSPASEWHDHFLNRTPITWLEFTVAPRKHIFSGVVSFKKKEPRAG
KFLGNDELSTYDGHDFQKLVDAIRDEFTVAPRKHIFSGVVSFKKKEPRAG
PRUTLQPSQGGSLSVVRPQFFNNAGNIRNQAAPRVAASQOQPAKREHSGK
PVYFCGDPGHYADCKPRVKVPTQSNSTVPASKARVNVHAAEAQADPVLIG
TFVNLVPAFMLFDSGATHSFLSMFAGNHGMEVDLRLPLMSTVSTQNALQSRS
VTEIOGVPLANLILLESKDLDVILGMDMLARHGVIDCANRKTILTSDGRVTVH
ALSSLSRLNQITLIEPIVREYDVPFDDLPMPKPRDEFRIDLVPGTPIHKRP
YMAANELAEVRKQDDLLQGVIRPSSSPWGAPEVFEKDHDTQRCMVDYRALNDV
IKNKYLLPRIDDLFDQKATVFSKIDLSRGYHQLRIKEEDIPTKATFTYRFPKCTV
HSYGLTNAPAFFNLMNVMEYLDKFFVVFIDILITSRTKEEHEHLALEKFLGA
MQLYAKFSKCEFWLSEVFLGHVISGAVDPSNVSTVMKQPTVSRIGFLGIA
GYVKEECBQSFQELNRLISAPILIPDKPGFOVYCDASKLGLCVLMQDKVVA
YASQRLRPHEKNYPTHDLELAAYVHALKIMRHYLFGRTMYTDHLSKLYIFQPDNL
MRORWLELIDKMGHHPKAVNVAADSLRGYCNCTEGQRLPELCEKFEFLDL
GIVSGFVALEAKPTLIDQVREAQINDPDIQIKNNMRGKAIGFLEDEQGTWVLCG
RSCVPNKDLKDALKEADHTLYSIHPGSTMYQDLKERFWASMKREIAEYAVCDV
CQVKAHEKQAGLQPLKIPKWEKEIGMDFTGLPRTSSGSHDSIWIIVDLRTKVAH
FIPVKVTSYSGSLAEALYMARIVCLHGVPKKIVSDRGSOFTSNFWKKLOEBSKLNFS
TAYHPQDQOTERVNQILEMDLACALDFGSGWDLPLVAEFSYNNYSQASLQMAPYE
ALYRQKRTPLMDQIGERQVFGTDLRBAEKVKLIQERLVAQRHKSIAQNRARD
LSPFEGDYVIRVTPFRFGVHTKGLAPFRFVGYKIVSRGGEVAYQLELPQSLAGV
HNVFHVSQLKCLRPVTEANLQIEVQEDLTVEKPIRILETNRRTRNRVIRFCVKV
OMSNHGBESTWTERDELKSAHPHLPASSSESGRDSV"
complement(28587..30710)
/gene="OSJNB0008A05.4"
/note="Contains similarity to gypsy-type retrotransposon
protein"
complement(join(28587..29387,30369..30457,30653..30710))
/gene="OSJNB0008A05.4"
/codon_start=1
```



```
/product="Putative gypsy-type retrotransposon protein"
/protein_id="AA04151.1"
/db_xref="GI:22655734"
/translation="MVRGGLCYARGIVTTPFRACGLEWNIPLGLEIELYSRDGCEWL
GNSNMRYKPEYLFVGGFQIEUTMSFVIGFGTAPYAIQIPHRHBEKCRVKVTLN
SLNEDIPVMEAGRNTHACQEVARIARELDRIYDQADTETRYRHPROSGSDR
GSYLETENDATRHIVLWAMDEGRAETVLAADQREDNRGKI CKLEBDKVDRLD
KELATLKEAPQKARIILTKARKALFVRYVQLAPKRVVEKEVAPTLDAPVLIIS
DEEBESKRNHASKIEWAQDDEDEPNESIDLRCSEELE"
complement(30001..32200)
/note="Similar to truncator transposon"
/gene="OSJNB0008A05.5"
complement(32582..33127)
/note="Hypothetical protein"
complement(32582..33127)
/gene="OSJNB0008A05.5"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AA04152.1"
/db_xref="GI:22655735"
/translation="MGPIVNPFSLSFSAHLPPSPGCGAATFRGSCTRLRTRTST
SYTRACSLVPLGAPRGRVVAHRLPGALRLDIPGHPARGDDERMSPASNNLA
QSVNCCURCASSGAMRTWPSATSRGCCATCTRSPTACTSSSTASTRRSSPTALG
RAFREDLLKGANLVLTPT"
complement(34658..36447)
/gene="OSJNB0008A05.6"
/note="Hypothetical protein"
complement(34658..36447)
/gene="OSJNB0008A05.6"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AA04153.1"
/db_xref="GI:22655736"
/translation="MDINMYTGHCKLSRRSVQPKRIATFTFLFIITRONLAPNAGL
HRLSISGGDQTDWQNLDTLRLRPLLRHPKIMNLRGHWKMSDANFNNTIIS
AFI"
complement(38823..38999)
/note="Similar to Explorer_Os1 MITB-like element"
complement(40530..41279)
/gene="OSJNB0008A05.7"
/note="Hypothetical protein"
complement(40530..41279)
/gene="OSJNB0008A05.7"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AA04154.1"
/db_xref="GI:22655737"
/translation="MTPPPSQTEISQDNLIPATMSLNDEQWAEIGNVMVQQLQEACLK
TPSIAQGIIVQKTFKRTSCGREGSCTDIKVEQDEEMQKVLTKFNQIAIDHVDISA
VNHALINOSVLVNTLANLIKSVDPGSAEHERGSLVLPGGFNPYRELKTSIGGGR
FQDNVMYRSKQPTSGVFGSGNTWPEKALILKSLVPGSTQPSATEFIARRPPVIP
QPTQPSGQPIGSHITRPQYCIQPRFSF"
3103..443154
/note="Simple Sequence Repeat (AT)n"
43830..44327
/gene="OSJNB0008A05.8"
/note="Unknown protein"
43830..44327
/gene="OSJNB0008A05.8"
/codon_start=1
/product="Unknown protein"
/protein_id="AA04155.1"
/db_xref="GI:22655738"
/translation="WAEENKSGSLMASGKVAETMSVFOQKSGVEGVYDKEVAGAA
ADLLHASTYIGKLDKPVQYIDKAEYLKDFSSGSGAARPPFPAAGDRAAPKPAEE
PPKEPAPAEKKGKPPSESGFLDDVMKGASSLMKKGGGSESGAGGAGGLFKMA
QGFMK"
46561..47056
```

```
repeat_region
Query Match 2.7%; Score 48; DB 8; Length 124366;
Best Local Similarity 100.0%; Pred. No. 4.9e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 30 AAATAAATAAATAATCCAAATTTTTTAAATAGACGAGTGGTCAACA 77
|||||
Db 68083 AAATAAATAAATAATCCAAATTTTTTAAATAGACGAGTGGTCAACA 68130
|||||

AC091749 150465 bp DNA linear PLN 13-NOV-2001
Oryza sativa chromosome 10 BAC OSJNB0008A05 genomic sequence,
complete sequence.
AC091749
AC091749.8 GI:16445110
HTG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 150465)
Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Moffat, K.S., Hill, J.N.,
Gansberger, K., Brenner, M., Burgess, S., Hance, M., Shvartsbeyn, M.,
Tsitrin, T., Riggs, F., Hsiao, J., Zismann, V., Blunt, S., Pai, G.,
VanAken, S.E., Utterback, T.R., Feldblyum, T.V., Kalb, E.M.,
Quackenbush, J., Salzberg, S.L., White, O. and Fraser, C.M.
Oryza sativa chromosome 10 BAC OSJNB0008A05 genomic sequence
Unpublished
2 (bases 1 to 150465)
Buell, R.
Direct Submission
Submitted (30-MAY-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 150465)
Buell, R.
Direct Submission
Submitted (26-OCT-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
4 (bases 1 to 150465)
Buell, R.
Direct Submission
Submitted (13-NOV-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
On Oct 26, 2001 this sequence version replaced gi:15808580.
Address all correspondence to:rice@tigr.org

BAC clone OSJNB0008A05 is from Oryza sativa chromosome 10
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
gscan and Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky,
http://genemark.biology.gatech.edu/Genemark/), and GeneSplicer
(Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
```

This BAC overlaps with rice BAC OSJNBa0073L20.

Location/Qualifiers

1..150465
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"

FEATURES

source

AB017076	Oryza sativa (japonica cultivar-group)	300029 bp	DNA	linear	PLN 06-JUN-2003	mrna
LOCUS	Oryza sativa (japonica cultivar-group)					
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 10, section 30 of 77 of the complete sequence.					
ACCESSION	AB017076	AE016959				CDS
VERSION	AB017076.1	GI:31431167				
KEYWORDS						
SOURCE	Oryza sativa (japonica cultivar-group)					
ORGANISM	Oryza sativa (japonica cultivar-group)					
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.					
REFERENCE	1 (bases 1 to 300029)					
AUTHORS	The Rice Chromosome 10 Sequencing Consortium					
CONSTRM	In-depth view of structure, activity, and evolution of rice chromosome 10					
TITLE	Science 300, 1566-1569 (2003)					
JOURNAL	2 (bases 1 to 300029)					
REFERENCE	Buell, C.R., Wing, R.A., McCombie, W.R., Messing, J. and Yuan, Q.					
AUTHORS	Direct Submission					
TITLE	Submitted (05-MAY-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA					gene
JOURNAL	This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted as tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).					mrna
FEATURES	Location/Qualifiers					CDS
source	1..300029					
	/organism="Oryza sativa (japonica cultivar-group)"					
	/mol_type="genomic DNA"					
	/cultivar="Nipponbare"					
	/db_xref="taxon:39947"					
	/chromosome="10"					
misc_feature	<1..90733					
	/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSUNBa0093109 (GB:AC090486)."					
gene	<1538..>2977					
	/locus_tag="OSUNBa0093109.20"					
	/note="hypothetical protein with F box domain"					
mrna	<1538..>2977					
	/locus_tag="OSUNBa0093109.20"					
CDS	1538..2977					
	/locus_tag="OSUNBa0093109.20"					
	/codon_start=1					
	/product="hypothetical protein"					
	/protein_id="AAP52983.1"					
	/db_xref="GI:31431168"					
	/translation="MSSPPPPPPCAACDGMWLSVSTVSGESNQKRLKRGGGGG GGGVGGGAVDDCGPHDEVLVLLVFAECSLETDLLVRCATCRWRRLVAGDAEYICR RKPPRRYVYCALAVGVQRRQENSSSSGAPPFRFPLPSYSSRFAAGGEGEDRKVFD SGLNSRLTASRGLLVLELRSSRAAAVRLVCPNMTGDMATLPILAKDRPGHYA CALITFDHDEGAPDRLGFVHDPAARLLVAVKRNFTACRSYMSDTKAWDAEGKLSGA KIGRRLEMTGAVVRGVSFWLLKNLLFVRLGALKATETFPFKWCSKLCFCYGP VQRQLAVTPDGRVTHVQVDRHVTNNINVRSHDGGTPTWECNARVDELNRV LPMANVRVCLRGVRSGLVFLAIGTDLYNQOPDLALYALDMDKKEARKVAAPPGHC RRLSSFFGYEMDRVAYLASLSGSESTAS"					
	complement (<11385..>17472)					
gene	/locus_tag="OSUNBa0093109.21"					
	/note="contains pfam domain pkinase protein kinase domain e value 7 4e 47; contains pfam domain DUF26 Domain of unknown function DUF26 e value 6 8e 19"					
complement	(join(<11385..11705,11746..11962,12443..12680,13029,13239,13320..13444,15245..15364,16641..17472))					
locus_tag="OSUNBa0093109.21"						
complement	(join(<11385..11705,11746..11962,12443..12680,13029,13239,13320..13444,15245..15364,16641..17472))					
locus_tag="OSUNBa0093109.21"						
codon_start=1						
product="putative serine/threonine kinase"						
protein_id="AAP52984.1"						
db_xref="GI:31431169"						
translation="MATRVLLSSKVVILALVLALESFLPRHGRALRLMWSNNGSAYA ANTYDYNVHSILATLTARTNTTGFATATTGRGTDEAMGLALCRGDTDRVGCASC LAAPAVAFNECRGDMVTYVFYRCARFYSYVDFTRPDNTEVLIGSPSADRITADAG HFDAVLADLAGALDAAAYNSTLYAAGVMTSGDGMSTTDMVNIYGVVQCTPDQA AAACRACLEALRVDPMPKVFAGMGRFNACNLYETFEVFDGDPSPVKLVAPPVPE DKKRSPENAAVVGALVILSLFIYLLWKLOAKOYAADKNVDSGSLFLDLA IIRKATANAENKLGCHGFAGYKGFLEDSIIVAEIARLDRTSGQLRQLNELLIYA KLHNNLAKLVGVCIKGDEKLIVFEFLPDKRQKREKSWETRIYIHGTA RGLLYHEDSQIKIHRDLKASNVLLSNMNPKISDFGLARICSGTKTITSQVVTG LGYMAPEYVLGHLSVKVDVYSGILVLEIVTGRNTDVFDADESSNLLSYRVIDQ LFYDPLKAMAFMLKVDHWQKGIPLTITDILLISGSRGLQDMELLKCVHGLLCVQ ENPADRPTMLSVLVYMLQVDVTNPAAPSKPAFTFANVRNTTSSSPSAAALSANEVSI SEFHR"						
complement	(<24816..>28021)					
locus_tag="OSUNBa0093109.22"						
note="Similar to PUTATIVE SERINE THREONINE KINASE RECEPTOR PRECURSOR (S RECEPTOR KINASE) (SRK) E value 8e 85 Identities 160 849 (of subject); contains pfam domain DUF26 Domain of unknown function DUF26 e value 2 8e 22"						
complement	(join(<24816..25133,25231..25384,25598..26006,26216,26305..26429,26847..26966,27145..>28021))					
locus_tag="OSUNBa0093109.22"						
complement	(join(<24816..25133,25231..25384,25598..26006,26216,26305..26429,26847..26966,27145..>28021))					
locus_tag="OSUNBa0093109.22"						
codon_start=1						
product="putative serine/threonine kinase"						
protein_id="AAP52985.1"						
db_xref="GI:31431170"						
translation="MAALVLMWPTLLPLVLLANLLPLRHRGALRLMWSNNGSSYA ANTYHSNVRAVLTAISAITNSRTARFATASAGCGADAVMGLACRGDTRDVGASC LAAPAVAFNECRGDMVTYVFYRCARFYSYVDFTRPDNTEVLIGSPSADRITADAG HFDAVLADLAGALDAAAYNSTLYAAGVMTSGDGMSTTDMVNIYGVVQCTPDQA AAACRACLETLRVDPMPKVFAGMGRFNACNLYETFEVFDGDPVKLVAPPVVPV DKKRSPENAAVVGALVILSLFIYWKKLOAKOYAADKNVDSGSLFLDLA IIRKATANAENKLGCHGFAGYKGFLEDSIIVAEIARLDRTSGQLRQLNELLIYA KLHNNLAKLVGVCIKGDEKLIVFEFLPDKRQKREKSWETRIYIHGTA RGLLYHEDSQIKIHRDLKASNVLLSNMNPKISDFGLARICSGTKTITSQVVTG LGYMAPEYVLGHLSVKVDVYSGILVLEIVTGRNTDVFDADESSNLLSYRVIDQ LFYDPLKAMAFMLKVDHWQKGIPLTITDILLISGSRGLQDMELLKCVHGLLCVQ ENPADRPTMLSVLVYMLQVDVTNPAAPSKPAFTFANVRNTTSSSPSAAALSANEVSI SEFHR"						
complement	(<24816..>28021)					
locus_tag="OSUNBa0093109.23"						
note="Similar to (AC021891) putative transposable element (Oryza sativa) E value 0 0 Identities 367 817 (of subject); contains pfam domain Transposase_21 Transposase family tnp2 e value 2 9e 31"						
join	(<29419..29431,29533..30336,30684..31133,31380..31469,31554..31723,31783,32347,32404..>32693)					
locus_tag="OSUNBa0093109.23"						
join	(<29419..29431,29533..30336,30684..31133,31380..31469,31554..31723,31783,32347,32404..>32693)					
locus_tag="OSUNBa0093109.23"						
codon_start=1						
product="putative TNP2 transposase"						
protein_id="AAP52986.1"						
db_xref="GI:31431171"						
translation="MDQSTIDRSWMYAKNLKHSYRKGVINFNMAAEEEDRKRNSD YMCPCACKDNEMFDSGEDVHGHOIQRGFMKGVTCTWKGHEQELSGASADRSAYN QEDDELMLIPSLGDENVADHDLQDMCLDVEDPAQNERDGMKFRSLVDSGTPL YVGKAKHTKLSVTLDLMLKLSKNGWTDKSFDDLGLKAMLPVNTLPTTYEAKOV LCPGLGVRRIHACPNDCILYHKQYADLDADYCKASVKRKKSGDNKRGKGPQO PCNDIDVFLFEPIDDLLEILWKEGMEWTWYQGENFKLRVLLFCTINDYFALGNLSQO						

```
IKKKACSDCKEHTSRSLKSRKRWVYMGHRSWLPRLHRAFRKKKIFNGKRELOLAPK
DLSGDEVHNVKVDISNEFGKRRKSKKACHKACHTSLSKDEKIAMLSCLKDIKVPKSGSA
RITAILGILPVKQITLQRCAPENALGQKVIDPDLGLOTLVNTLCHLEHFFPVH
LVGSIIEGTYTTEKALIFCVNMSDADPIGVPSRHGRSLSGVGTIGKRLSPQASV
AQHAYVLQHMAGSGSVFSEHLAKIRDENIGRSDAWISREHNSRFNFKVRVTMSYD
VNEIIVLQSGMSWTDVWQGYDINGTYFYTKQDQIGVQNSGVRIIDAFQDQVGN
TYDRIFEIWEVNDYSDKEGFTLVLDLSKGVADPEFLAKQVKQIFIKDPSNKKMHI
VRDQKCRIVGVNVDDEEYNNHNLHVRPILDLDDDDPQEPVAYASDTEGTLT
<34963..>38789
/locus_tag="OSJNBa0093109.24"
/notes="hypothetical protein"
join(<34963..35178,35671..36402,36484..36927,37014..37061,
37102..37901,37980..38150,38233..38521,38607..38645,
38751..>38789)
/locus_tag="OSJNBa0093109.24"
join(34963..35178,35671..36402,36484..36927,37014..37061,
37102..37901,37980..38150,38233..38521,38607..38645,
38751..38789)
/locus_tag="OSJNBa0093109.24"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP52987.1"
/db_xref="GI:314311172"
/translation="MSELIIFVAYSSPTDAAHPPTCTPTANSIAEPAAATEPAAST
EPGTASTDRRPPSPPLSPGPRQPMQSGPEPMTTSSGSRARTPTGKIKPK
ERFRTAVDAVLPTRKILCRPSICGVIGROKFSILODDIKLVPAAKDIAMLFM
KESFDYPAEHEDRLRAAFKVMGNWKNFKTLGVEFYVNPANPDPREKFPWITEQVM
EEHAKSTPESKASSAYRLLTQNHQPHRLGTGTGAGKEEWQREDEAESNTPL
VFGDIPHRANRWARYQKNDGTIFMPNPEDQRYKALELVAEQASQACSCQCE
DDLTLKALGNHGRHTQGVSSVYKYPFDYAWQYKRTSNKAEDARLEAKIRS
IRAEITSEFDEKLESMKATRKETREQQNPQAAATTHELGASPTQKHSKASTELA
ENPTSVDASVDHTLEPTSCITLVMPFMQECERNSGDCACAKOVDSVKPEYELP
PLKYPNDVLSQMDGEBRTEADSKAIEHPQMKSRKAHSPQRTILDKPEAKRQ
DRAQSDVQLDYMDFGEBRTEADSKAIEHPQMKSRKAHSPQRTILDKPEAKRQ
GRGKQVQSLAPKDLQDQBEETKGEKVKYALQRFGLMPLGVDDVLAAMGTA
CKDILHAYMEKSNKNTKATDILGEGKPFPLGNTYIVIDFKDLPLRLRAVDTC
LLKCYLSLQWCKNTPEVAFDLPQVTVTNLQNDQGMVNTYDFTLWTRDKEITYM
CAYNOYAHWILLVTPKWSCHYLSNRIDKNAYDWTPIOLAIDEAWAQYVQRGSLRS
GHTLIHKDPKPMKQIQDQCGFHVCHNMLLYREKVKTLAEFGTITKSLPTSFBEA
MNTYIATVM"
complement(<43954..>49277)
/locus_tag="OSJNBa0093109.25"
/notes="Similar to PHOSPHATIDATE CYTIDYLTRANSFERASE (CDP
DIGLYCERIDE SYNTHETASE) E value 7e 50 Identities 96 424
(of subject); contains pfam domain Cytidylyltrans
Phosphatidate cytidylyltransferase e value 2 7e 41"
complement(join(<43954..44031,44113..44142,44586..44642,
44713..44841,44947..45021,45155..45303,45452..45506,
48635..48701,48789..49042,49161..>49277))
/locus_tag="OSJNBa0093109.25"
/codon_start=1
complement(join(43954..44031,44113..44142,44586..44642,
44713..44841,44947..45021,45155..45303,45452..45506,
48635..48701,48789..49042,49161..>49277))
/locus_tag="OSJNBa0093109.25"
/codon_start=1

Query Match 2.78; Score 48; DB 8; Length 300029;
Best Local Similarity 100.0%; Pred. No. 4.3e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 AAATAAATAATAATCCAAATTTTTTTTAAATAAGACGAGTGCTCAACAA 77
|||||
Db 262872 AAATAAATAATAATCCAAATTTTTTTTAAATAAGACGAGTGCTCAACAA 262919
|||||

RESULT 14
AX656219/c
LOCUS AX656219 2000 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 6089 from Patent WO03000898..
ACCESSION AX656219
VERSION AX656219.1 GI:29159033
KEYWORDS
SOURCE Oryza sativa
```

```
ORGANISM Oryza sativa
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 6089 03-JAN-2003;
SYNGENTA PARTICIPATIONS AG (CH)
FEATURES
source 1. 2000
Location/Qualifiers
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Query Match 2.5%; Score 45; DB 6; Length 2000;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 CAAATTTTTTAAATACGAGTGCTCAACACAGTACAGTAAAAA 90
|||||
Db 294 CAAATTTTTTAAATACGAGTGCTCAACACAGTACAGTAAAAA 250
|||||

RESULT 15
AF254558/c
LOCUS AF254558 4817 bp DNA linear PLN 11-MAR-2001
DEFINITION Oryza sativa NAC6 (NAC6) gene, complete cds.
ACCESSION AF254558
VERSION AF254558.1 GI:13272280
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 4817)
AUTHORS Yoon, U.-H., Hahn, J.-H. and Eun, M.-Y.
TITLE Molecular cloning of NAC6 gene in rice
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4817)
AUTHORS Yoon, U.-H., Hahn, J.-H. and Eun, M.-Y.
TITLE Direct Submission
JOURNAL Institute of Agricultural Science and Technology, National
Suwon 441-707, Korea
FEATURES
source 1. 4817
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="48101"
727..>4586
/gene="NAC6"
727..766
/gene="NAC6"
join(<2151..2316,2416..2690,3552..>4586)
/gene="NAC6"
/product="NAC6"
join(2151..2316,2416..2690,3552..4022)
/gene="NAC6"
/codon_start=1
/product="NAC6"
/protein_id="AAK17067.1"
/db_xref="GI:13272281"
/translatation="MSGGQDLQPPGFRPHPTDEELVMHYLCRCAGLPIAVPIIAEI
DLYKFDPPWOLPRMALYGEKEWYFSPDRKYPNGSRPNRAAGSGYWKATGADKPVGSP
KPAVIAIKALVYAGKAPKGEKTNWIMHEYRLADVDRSARKNSLRLLDDWLCLRIYNNK
GGLEKPPAAVAAGVSSGGGVQKPMGVNNAVSPPEQKPVVAGPAPDLAAYYD
```

RPSDMPRLHADSSCSEQVLSPEFACEVQSOPKISEWERTFATVGPINPAAASILDPA
 SGGGLGGGGSDPLLQDILMYWGKPF"
 4581. .4586
 /gene="NAC6"

polyA_signal

ORIGIN

Query Match 2.5%; Score 45; DB 8; Length 4817;
 Best Local Similarity 100.0%; Pred. No. 3.1e-12;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 CAAATTTTAAAGACGAGTGGTCAAAACAGTCAAGTAAAAA 90
 |||||
 Db 446 CAAATTTTAAAGACGAGTGGTCAAAACAGTCAAGTAAAAA 402
 |||||

Search completed: October 2, 2004, 13:18:15
 Job time : 11071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 07:17:19 ; Search time 6888 Seconds
(without alignments)
7682.315 Million cell updates/sec

Title: US-10-009-570-1
Perfect score: 1772
Sequence: 1 actttagataataagaag.....ctcttctcttcagtgaag 1772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	50	2.8	628	28	AQ686477 nbxb0072B
2	38	2.1	748	28	AQ686991 nbxb0032J
C 3	30	1.7	569	28	AQ272200 nbxb0027L
4	30	1.7	629	14	CA754682 BR0300040

5	29	1.6	537	28	AQ510840
C 6	29	1.6	537	28	AQ510840
C 7	28	1.6	331	29	CG408907
C 8	28	1.6	409	28	AQ331226
9	28	1.6	452	28	AQ860090
C 10	28	1.6	466	28	AQ863506
C 11	28	1.6	623	28	AQ258100
C 12	28	1.6	652	12	BM420232
13	28	1.6	694	28	BZ692356
14	28	1.6	783	29	CC692859
C 15	28	1.6	787	29	CC692657
C 16	28	1.6	833	29	CG164014
C 17	28	1.6	856	28	AQ258985
C 18	28	1.6	887	29	CC704404
19	27	1.5	525	28	BH776225
20	27	1.5	728	14	CB663138
C 21	27	1.5	898	28	AQ914858
C 22	26	1.5	162	29	AG025502
23	26	1.5	296	29	CC833513
24	26	1.5	475	12	BI323424
25	26	1.5	493	12	BI450665
C 26	26	1.5	660	28	AQ287509
27	26	1.5	663	28	AQ872280
C 28	26	1.5	670	28	BZ786579
C 29	26	1.5	677	28	CC400206
C 30	26	1.5	708	29	CG371013
C 31	26	1.5	712	29	CG916678
32	26	1.5	761	28	AQ158754
C 33	26	1.5	798	28	AZ126642
C 34	26	1.5	810	29	CC760816
35	26	1.5	841	29	CG734891
36	26	1.5	862	28	BZ962486
37	26	1.5	898	28	AQ914858
38	26	1.5	970	28	CC370662
C 40	25	1.4	531	28	CC037205
C 41	25	1.4	566	28	BH788889
C 42	25	1.4	566	28	CC019996
C 43	25	1.4	637	28	AQ510347
C 44	25	1.4	651	29	CG286906
45	25	1.4	678	29	CG007640

ALIGNMENTS

RESULT 1
AQ686477/c
LOCUS
DEFINITION

AQ686477 628 bp DNA linear GSS 01-JUL-1999
nbxb0072B19r CUGI Rice BAC Library Oryza sativa (japonica
cultivar-group) genomic clone nbxb0072B19r, genomic survey
sequence.

ACCESSION
AQ686477
VERSION
KEYWORDS
SOURCE
ORGANISM

AQ686477.1 GI:5327645

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 628)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends

High quality sequence stop: 384.

FEATURES
source

```
1..628
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0072B19r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library"
/note="Vector: pBeloBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
```

```
242
```

```
7 GATATAAGTAAGTCACAGAAAATAATAATTCCTCAATTTT 56
485 GATATAAGTAAGTCACAGAAAATAATAATTCCTCAATTTT 436
```

ORIGIN

```
Query Match 2.8%; Score 50; DB 28; Length 628;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 7 GATATAAGTAAGTCACAGAAAATAATAATTCCTCAATTTT 56
Db 485 GATATAAGTAAGTCACAGAAAATAATAATTCCTCAATTTT 436
```

RESULT 2

```
AQ868991 748 bp DNA linear GSS 03-NOV-1999
LOCUS nbxb0032J14r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone nbxb0032J14r, genomic survey sequence.
```

```
ACCESSION AQ868991
VERSION AQ868991.1 GI:6219442
```

KEYWORDS

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 748)

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Wing RA

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: GGAAACAGCTATGACCATG

Class: BAC ends

High quality sequence start: 34

High quality sequence stop: 416.

Location/Qualifiers

FEATURES

source

```
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0032J14r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/note="Vector: pBeloBAC11; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

```

ORIGIN

```
Query Match 2.1%; Score 38; DB 28; Length 748;
```

```
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 29 AAAATAATAATAATTCCTCAATTTTAAATAAGACGA 66
425 AAAATAATAATAATTCCTCAATTTTAAATAAGACGA 462
```

RESULT 3

```
AQ272200/c
```

LOCUS

DEFINITION

AQ272200 569 bp DNA linear GSS 03-NOV-1998
nbxb0027L15f CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0027L15f, genomic survey sequence.

```
ACCESSION AQ272200
```

```
VERSION AQ272200.1 GI:3825515
```

KEYWORDS

SOURCE

ORGANISM Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 569)

AUTHORS Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Rice Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Wing RA

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: TAATACGACTCACTATAGGG

Class: BAC ends

High quality sequence stop: 316.

Location/Qualifiers

1..569

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

```

/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0027L15f"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library"
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN
Query Match      1.7%; Score 30; DB 28; Length 569;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAAATAAATAATATCCAAATTTTTTAA 58
    |||||||
Db 496 AAAATAAATAATATCCAAATTTTTTAA 467

RESULT 4
CA754682
LOCUS
DEFINITION
BR030004000_PLATE_E02_13_017.ab1 OA Oryza sativa (japonica cultivar-group) cDNA clone BR030004000_PLATE_E02_13_017.ab1 similar to No protein alignment, mRNA sequence.
CA754682
ACCESSION
VERSION
KEYWORDS
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 629)
REFERENCE
AUTHORS
Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrera,H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.
TITLE
Functional Genomics of Plant Stress Tolerance
JOURNAL
Unpublished (2000)
COMMENT
Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
Location/Qualifiers
1. .629
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="BR030004000_PLATE_E02_13_017.ab1"
/tissue_type="roots"
/dev_stage="3-4 weeks"
/clone_lib="OA"

FEATURES
source
1. .629
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/db_xref="taxon:39947"
/clone="nbxb0094A05r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library"
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

```

```

/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbxb0027L15f"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library"
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN
Query Match      1.7%; Score 30; DB 14; Length 629;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AAAATTTTAAAGACGAGTGTGTCAAAC 76
    |||||||
Db 230 AAAATTTTAAAGACGAGTGTGTCAAAC 259

RESULT 5
AQ510840
LOCUS
DEFINITION
nxb0094A05r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0094A05r, genomic survey sequence.
AQ510840
ACCESSION
VERSION
KEYWORDS
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 537)
REFERENCE
AUTHORS
Wing,R.A. and Dean,R.A.
TITLE
A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL
Unpublished (1998)
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 59
High quality sequence stop: 437.
Location/Qualifiers
1. .537
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/db_xref="taxon:39947"
/clone="nbxb0094A05r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library"
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

```

ORIGIN

Query Match 1.6%; Score 29; DB 28; Length 537;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 AATTTTATAAGACGAGTGGTCAAC 76
 |||||
 Db 508 AATTTTATAAGACGAGTGGTCAAC 536

RESULT 6
 LOCUS AQ510840/c
 DEFINITION nbx0094A05r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbx0094A05r, genomic survey sequence.

ACCESSION AQ510840
 VERSION AQ510840.1 GI:4732218
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 537)
 WING, R.A. and DEAN, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998)
 JOURNAL Contact: Wing RA
 COMMENT Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAAACAGCTATGACATG
 Class: BAC ends
 High quality sequence start: 59
 High quality sequence stop: 437.

FEATURES
 source
 1..537
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="nbx0094A05r"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library"
 /note="Vector: pBelobAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 1.6%; Score 29; DB 28; Length 537;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 AAAATAATAATAATTCCTCAAAATTTTT 56
 |||||
 Db 445 AAAATAATAATAATTCCTCAAAATTTTT 417

RESULT 7
 LOCUS CG408907/c
 DEFINITION Ds601 Ds insertion lines Oryza sativa (japonica cultivar-group) genomic, genomic survey sequence.

ACCESSION CG408907
 VERSION CG408907.1 GI:34430272
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 331)
 KIM, C.M., PIAO, H.L., PARK, S.J., CHON, N.S., JE, B.I., SUN, B., PARK, S.H., PARK, J.Y., LEE, E.J., KIM, M.J., LEE, J.J., NAM, M.H., SUN, M.Y. and HAN, C.D.
 Rapid, large-scale generation of Ds transposon lines and analysis of Ds loci in rice Unpublished (2003)
 JOURNAL Contact: Chang-deok Han
 COMMENT Division of Applied Life Science, PMBBRC
 Gyeongsang National University
 Gajwa-dong 900, Jinju 660-701, South Korea

Tel: +82 55 751 6029
 Fax: +82 55 759 9363
 Email: cdhan@nongae.gsnu.ac.kr
 Location: chromosome 4 clone OSJNBb0034G17
 Class: transposon-tagged.

FEATURES
 source
 1..331
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Dongjin"
 /db_xref="taxon:39947"
 /clone_lib="Ds insertion lines"

ORIGIN
 Query Match 1.6%; Score 28; DB 29; Length 331;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AATAAGACGAGTGGTCAACAGTACAAG 84
 |||||
 Db 220 AATAAGACGAGTGGTCAACAGTACAAG 193

RESULT 8
 LOCUS AQ331226/c
 DEFINITION nbx0049E15r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbx0049E15r, genomic survey sequence.

ACCESSION AQ331226
 VERSION AQ331226.2 GI:6591316
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 409)
 WING, R.A. and DEAN, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998)
 JOURNAL Contact: Wing RA
 COMMENT On Dec 15, 1999 this sequence version replaced gi:4123076.

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288

Fax: 864 656 4293
Email: twing@clemson.edu
Seq primer: GGAACACGCTATGACCATG
Class: BAC ends
High quality sequence stop: 1.

FEATURES
source Location/Qualifiers

1. 409
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"

/strain="Japonica"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="nxb0049e15r"

/tissue_type="Leaf"

/lab_host="E. coli DH10B"

/clone_lib="CUGI Rice BAC Library"

/notes="Vector: pBelobAC11; Site 1: HindIII; Site 2:

HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN

Query Match 1.6%; Score 28; DB 28; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 ATTTTAAATACAGAGTGTCACAC 76
|||||
Db 287 ATTTTAAATACAGAGTGTCACAC 260

RESULT 9

AQ860090

LOCUS

DEFINITION

AQ860090 452 bp DNA linear GSS 03-NOV-1999
nbeb0014121f CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone nbeb0014121f, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 452)

Wing, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Rice Genome

Unpublished (1998)

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: twing@clemson.edu

Seq primer: TAATACGACTCATATAGGG

Class: BAC ends

High quality sequence start: 31

High quality sequence stop: 404.

FEATURES

source

1. 452

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/strain="Japonica"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="nbeb0014121f"

/tissue_type="Leaf"

/lab_host="E. coli DH10B"

/clone_lib="CUGI Rice BAC Library (EcoRI)"

/notes="Vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI;

Rice is the most important food crop in the world. Half of

the world population, especially those inhabiting highly

populated areas of the humid tropics and subtropics, rely

on rice as their primary source of carbohydrate.

Monocotyledonous rice is a diploid plant (2n=24) with a

haploid genome equivalent of 431 Mbp (Arumuganathan and

Earle, 1991). The relatively small genome of rice, three

times larger than that of Arabidopsis, makes it suitable

for genomic studies. In order to facilitate positional

cloning, physical mapping and genome sequencing of rice,

we have constructed a BAC library from Oryza sativa,

Nipponbare variety using EcoRI as the cloning enzyme. The

library contains 55,296 clones with an average insert size

of 121 Kb providing approximately 15 haploid genome

equivalents. The deep coverage allows the isolation a

particular sequence with a probability of 99.9 %. Three

high density filters, each containing 18,432 clones

(doubly spotted), represent the whole library for colony

screening and can be requested from the Clemson University

BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN

Query Match 1.6%; Score 28; DB 28; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 TTTTAAATACAGAGTGTCACACA 77
|||||
Db 395 TTTTAAATACAGAGTGTCACACA 422

RESULT 10

AQ863506/c

LOCUS

DEFINITION

AQ863506 466 bp DNA linear GSS 03-NOV-1999
nbeb0021007f CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone nbeb0021007f, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 466)

Wing, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Rice Genome

Unpublished (1998)

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clmson.edu
 Seq primer: TAATACGACTCACTATAGG
 Class: BAC ends
 High quality sequence start: 28
 High quality sequence stop: 438.
 Location/Qualifiers
 1. 466
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="nbcb0021007f"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library (EcoRI)"
 /note="Vector: pBACindigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

FEATURES

source

Seq primer: GGAACACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 286.
 Location/Qualifiers
 1. 623
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="nbxb0019P11r"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library"
 /note="Vector: pBeloBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

FEATURES

source

ORIGIN

Query Match 1.6%; Score 28; DB 28; Length 623;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 AAATAAATAATTAATTCCAAATTTTTT 56
 |||||||||||||||||||||||||||||||
 Db 264 AAATAAATAATTAATTCCAAATTTTTT 237

QY 49 ATTTTAAATAAGACGAGTGTCAAAC 76
 |||||||||||||||||||||||||||||||
 Db 283 ATTTTAAATAAGACGAGTGTCAAAC 256

RESULT 11
 A0258100/c 623 bp DNA linear GSS 23-OCT-1998
 LOCUS nbxb0019P11r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbxb0019P11r, genomic survey sequence.
 DEFINITION A0258100
 VERSION A0258100.1 GI:3782582
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.
 REFERENCE 1 (bases 1 to 623)
 AUTHORS Wing, R.A. and Dean, R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu

RESULT 12
 BM420232

LOCUS BM420232

DEFINITION BM420232

ACCESSION BM420232

VERSION BM420232.1

KEYWORDS EST.

SOURCE Oryza sativa

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

REFERENCE 1 (bases 1 to 652)

AUTHORS Dong, H.T., Li, D.B., Zhuang, X.F., Dai, C.G., Sun, L.X., Pei, Y.X., Wu, H.F., Jiang, Y.X., Yu, F.C., Gao, Q.K. and Lou, Y.C.

TITLE A Gene Expression Screen in Oryza sativa

JOURNAL Unpublished (2001)

COMMENT Contact: Haitao Dong, Dehao Li
 Bioinformatics and Gene Network Research Group
 Zhejiang University
 Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
 Tel: 0086-571-86892051
 Fax: 0086-571-868961525
 Email: webmaster@estarray.org, URL: <http://www.estarray.org>
 Seq primer: M13 forward primer.

FEATURES

source

1. 652

/organism="Oryza sativa"

/mol_type="mRNA"
 /db_xref="taxon:4530"
 /clone="U004F12"
 /tissue_type="leaf"
 /dev_stage="Mature stage"
 /clone_lib="Oryza sativa mature leaf library induced by
 M.grisea"
 /note="Vector: pSport2"

ORIGIN

Query Match 1.6%; Score 28; DB 12; Length 652;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AATAAGACGAGTGTCAACAGTACCAAG 84
 ||||||||||||||||||||||||||||||||
 Db 542 AATAAGACGAGTGTCAACAGTACCAAG 569

RESULT 13
 BZ692356
 LOCUS
 DEFINITION SP_Ba0019009.f SP_Ba Sorghum propinquum genomic clone
 SP_Ba0019009 5', genomic survey sequence.
 ACCESSION BZ692356
 VERSION BZ692356.1 GI:28384160
 KEYWORDS GSS.
 SOURCE Sorghum propinquum
 ORGANISM Sorghum propinquum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 694)
 AUTHORS Wing, R., Yu, Y., Kim, H.R., Collura, K., Pries, G., Currie, J.,
 Soderlund, C., and Hatfield, J.
 TITLE Sequencing of Sorghum BAC ends.
 JOURNAL http://genome.arizona.edu/stc/sorghum
 COMMENT Unpublished (2003)
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: atc agc ggc cgc gat cc
 BACKWARD: gta aaa cga cgg cca gtc
 Plate: 0019 row: 0 column: 09
 Seq primer: atc agc ggc cgc gat cc
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1..694
 /organism="Sorghum propinquum"
 /mol_type="genomic DNA"
 /db_xref="taxon:132711"
 /clone="SP_Ba0019009"
 /clone_lib="SP_Ba"
 /note="Vector: pBel0BAC11; Site_1: HindIII; Site_2:
 HindIII; Paterson lab BAC library (HindIII)"

ORIGIN

Query Match 1.6%; Score 28; DB 28; Length 694;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 AATTTTTTTAATAAGACGAGTGTCAAA 75
 ||||||||||||||||||||||||||||||||
 Db 532 AATTTTTTTAATAAGACGAGTGTCAAA 559

RESULT 14
 BZ692356
 LOCUS
 DEFINITION SP_Ba0019009.f SP_Ba Sorghum propinquum genomic clone
 SP_Ba0019009 5', genomic survey sequence.
 ACCESSION BZ692356
 VERSION BZ692356.1 GI:28384160
 KEYWORDS GSS.
 SOURCE Sorghum propinquum
 ORGANISM Sorghum propinquum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 694)
 AUTHORS Wing, R., Yu, Y., Kim, H.R., Collura, K., Pries, G., Currie, J.,
 Soderlund, C., and Hatfield, J.
 TITLE Sequencing of Sorghum BAC ends.
 JOURNAL http://genome.arizona.edu/stc/sorghum
 COMMENT Unpublished (2003)
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: atc agc ggc cgc gat cc
 BACKWARD: gta aaa cga cgg cca gtc
 Plate: 0019 row: 0 column: 09
 Seq primer: atc agc ggc cgc gat cc
 Class: BAC ends.

CC692659
 LOCUS
 DEFINITION OGTAI20TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0389C15,
 genomic survey sequence.
 CC692659
 ACCESSION CC692659
 VERSION CC692659.1 GI:32097435
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 783)
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGTAI20TC
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source
 Location/Qualifiers
 1..783
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0389C15"
 /clone_lib="ZM_0.7_1.5_KB"
 /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 1.6%; Score 28; DB 29; Length 783;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 AATTTTTTTAATAAGACGAGTGTCAAA 75
 ||||||||||||||||||||||||||||||||
 Db 293 AATTTTTTTAATAAGACGAGTGTCAAA 320

RESULT 15
 CC692657/c
 LOCUS
 DEFINITION OGTAI20TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0389C15,
 genomic survey sequence.
 CC692657
 ACCESSION CC692657
 VERSION CC692657.1 GI:32097433
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 787)
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGTAI20TH
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208

RESULT 15
 CC692657/c
 LOCUS
 DEFINITION OGTAI20TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0389C15,
 genomic survey sequence.
 CC692657
 ACCESSION CC692657
 VERSION CC692657.1 GI:32097433
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 787)
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGTAI20TH
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208

RESULT 15
 CC692657/c
 LOCUS
 DEFINITION OGTAI20TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0389C15,
 genomic survey sequence.
 CC692657
 ACCESSION CC692657
 VERSION CC692657.1 GI:32097433
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 787)
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGTAI20TH
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208

Email: whitelaw@tigr.org
Seq primer: Tg
Class: sheared ends.

FEATURES
source Location/Qualifiers

1..787
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0389C15"
/clone_lib="ZM_0.7.1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 1.6%; Score 28; DB 29; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 48 AATTTTAAAGACGAGTGGTCAAA 75
Db 495 AATTTTAAAGACGAGTGGTCAAA 468

Search completed: October 2, 2004, 15:13:13
Job time : 6892 secs